Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JAN 1999		**00-00-1999 to 00-00-1999**

4. TITLE AND SUBTITLE
Literature Survey for Issues in Naval Decision Support: Phase II

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
HumanSystems Incorporated,111 Farquhar Street,Guelph, Ontario, Canada N1H 3N4 ,

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR’S ACRONYM(S)

11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**227**	

HUMANSYSTEMS
Incorporated

Humansystems

111 Farquhar Street

Guelph, Ontario

N1H 3N4

Tel: (519) 836-5911

Fax: (519) 836-1722

http://www.humansys.com

Literature Survey for Issues in Naval Decision Support: Phase II

PWGSC Contracts #W7711-7-7426/000/SRV
W7711-7-7404/01-SV
Call-up 7404-04

January 1999

DCIEM. No. CR 1999-014

DCIEM No.

Literature Survey for Issues in Naval Decision Support: Phase II

by:

D.J. Bryant, R.D.G. Webb

Human*systems*, Incorporated
111 Farquhar St., 2nd floor
Guelph, ON N1H 3N4

Project Manager: D.J. Bryant

(519) 836 5911

PWGSC Contracts #W7711-7-7426/001/SRV
W7711-7-7404/01-SV
Call-up 7404-04

On behalf of
DEPARTMENT OF NATIONAL DEFENCE

DCIEM Scientific Authority
Carol McCann
(416) 635-2190

January 1999

Table of Contents

Executive Summary

The Command Decision Support Working Group will formulate a Research and Development plan for supporting enhanced operational effectiveness for the Command Team in the HALIFAX class through improved situation awareness and decision support. To aid in the development of this plan, a literature search and review was completed in two phases. In Phase I, a database of articles relevant to the naval domain was created using EndNote software. Phase II addressed the further search of the database, the review of articles, and the compilation of conclusions regarding:

1. Results, recommendations, approaches, and guidelines from studies that are immediately usable by the Navy in the upgrade of the HALIFAX class;
2. Problems in the design, development, and implementation of military tactical or operational situation awareness and decision support that have hampered or precluded success of these systems ("lessons learned") and how these problems might be avoided;
3. Issues in naval team decision support that require further research and the implications of not addressing them.

This report documents the results of Phase II. The literature review serves as an overview of many different perspectives and an analysis of each perspective. Overall, the literature review is intended to provide both theoretical analysis and practical guidance. Consequently, each major subsection contains recommendations directly aimed at the HALIFAX class upgrade.

Overall, the report is divided into nine major sections.

Introduction. This section describes the purpose and scope of the project, context, assumptions, and limitations.

Theory. This section reviews issues and theoretical perspectives surrounding decision making and decision support for naval Command and Control (C2) and discusses the naval tactical environment (including definitional issues surrounding command and control and the anticipated emphasis on littoral warfare in the future), theories of decision making (focusing on the contrast between analytic and intuitive theories and their implications for decision support), situation awareness (Endsley's three-level theory), expertise and training (stages of expertise and experiential learning techniques), teams (theories of teamwork, team decision making, and team situation awareness), and human-computer interaction (cognitive fit between interface and user).

Empirical Results. This section outlines major empirical findings pertinent to the theoretical issues raised in the first section and discusses results pertaining to decision making (applicability of analytic and intuitive theories and factors affecting decision making), decision support systems (approaches, issues, and examples), situation awareness (errors and factors affecting situation awareness), expertise and training (evaluation of the stages of expertise and approaches to training), and human-computer interaction (display and input issues).

Summary of Theory and Empirical Results in the Operational Context. This section discusses the issues raised in the preceding two sections in relation to the operational context for the HALIFAX class. The purpose is to highlight issues and questions of particular relevance to the design of decision support.

Methodologies. This section discusses ways to further study decision making and related processes, including design processes for developing decision support systems (user-centred design and rapid prototyping), research tools for evaluating decision support (scenario-based and cellular automata

simulations, evaluation methods and issues), measures of effectiveness (classification and techniques), analysis techniques (cognitive task analysis and conceptual graphing), team research techniques (principles and methods), and human-computer interaction (design and evaluation methods).

Summary of Methodology in the Operational Context. This section reviews methodological issues as they are related to the operational context.

Guidelines. This section reviews guidelines developed for naval C2 systems, as well as general guidelines for Human-Computer Interaction (HCI).

Summary of Guidelines in the Operational Context. This section relates issues of guidelines to the operational context.

Conclusions. This section summarizes the major findings of the literature review and draws the following major conclusions:

1. Future literature reviews should consider relevant topics in the broader, non-naval literature on Situation Awareness (SA) and decision making, taking care to avoid extensive, unproductive literature searches by focusing on only highly related research.
2. The future naval environment will place greater demands on naval decision makers who will be required to act rapidly, under high risk, with uncertain information, in novel and unfamiliar situations.
3. Tactical decision makers primarily use intuitive decision making strategies (e.g., recognition and story generation), indicating a need for decision support to overcome the weaknesses of these strategies (i.e., no analysis of the quality of solutions, dependence on familiarity of situations).
4. Research has tended to ignore mission planning and other preparation activities that are part of C2 but these areas are crucial and may be supported by analytic theories of decision making.
5. Further research should be conducted to clarify the processes by which people develop SA, the specific requirements for SA support, and the relation of SA to decision making.
6. Naval C2 places a great emphasis on teamwork (within and between ships), indicating a need for further study of how teams coordinate, communicate, allocate tasks and resources, achieve team SA, and make decisions.
7. Researchers should follow user-centred design methods, involving users from the start to ensure that system requirements and design concepts will help users achieve their goals.
8. Researchers should make use of rapid prototyping methods to increase the effectiveness of user-centred design and provide users with concrete models of system features.
9. There is no single best method to evaluate system or human performance and researchers should adopt a broad research approach that makes use of experimental, observational, and simulation studies, conducted with high and low levels of fidelity to operational conditions (as demanded by empirical goals), and employing a range of measures of effectiveness.
10. Researchers should explore emerging technologies and techniques that can potentially enhance the Navy's ability to evaluate new C2 systems.
11. There is a need to develop team research techniques and methods to support research in teamwork and to integrate theories of teamwork with theories of decision making.
12. Survey existing Decision Support Systems (DSSs) to identify concepts, techniques, and HCI features to determine which are directly applicable to the HALIFAX class.
13. Survey scientific, engineering, and computer science literatures to identify practical guidelines for DSSs, C2 systems, and HCI.

1. Introduction

1.1 Background

The Navy has identified decision support as a focus of attention in the upgrade to the Canadian Patrol Frigate (CPF), which is anticipated for the 2005-2010 timeframe. The Command Decision Support (CDS) Working Group (WG) has been formed to determine the Research and Development (R&D) needed in this area. Specifically, the WG will formulate an R&D plan for supporting enhanced operational effectiveness for the Command Team (CT) in the HALIFAX class through improved situation awareness and decision support. To aid in the development of this R&D plan, a literature review was completed in two phases. In Phase I, a database of relevant articles was created using EndNote software. In Phase II, we selected approximately 200 articles and reviewed them in detail to form the basis of the literature review.

1.2 Purpose

In Phase I, we developed a search strategy and implemented it to search the National Technical Information Service (NTIS) and PsycInfo databases (Bryant, Webb, & McLean, 1998). The search yielded 6,143 articles of potential interest, from which we chose 122 as potentially relevant. We obtained a subset of these articles, reviewed them in more detail, and identified additional pertinent keywords. Phase II addressed the further search of the extracted database, the review of articles, and the compilation of conclusions as required below. The overall intent (Phases I and II) was to review the basic and applied research literature concerning the development, implementation, and use of naval decision support for single ship Command and Control (C2). The literature review identified:

1. Results, recommendations, approaches, and guidelines from studies that are immediately usable by the Navy in the upgrade of the HALIFAX class.
2. Problems in the design, development, and implementation of military tactical or operational decision support that have hampered or precluded success of these systems ("lessons learned") and how these problems might be avoided.
3. Issues in naval team decision support requiring further research and the implications of not addressing them.

The review concentrated first and foremost on findings that are in the context of single-ship naval decision support, although operations within task groups was also considered. It covered literature from North American, European, and Australian R&D.

1.3 Tasks

To achieve the purpose of Phase I of the literature review, the following tasks were performed:

1. Carry out an additional search for a) articles in the extracted database that are indicated by the revised set of keywords developed in Phase I, and b) articles that have been derived from the current set of relevant titles. *(Complete).*
2. Obtain the relevant articles identified by the searches conducted in Phases I and II. *(Complete).*
3. Review in detail the total set of relevant articles. *(Complete).*

4. Create a new EndNote bibliography consisting of complete references for all articles identified as of interest and including the notes made in the review. *(Complete).*
5. Co-ordinate with similar review of data on development, implementation, and use of shipboard tactical data fusion and resource management decision aids being undertaken at DREV, in terms of passing information, limiting duplication, and ensuring coverage. *(Incomplete).*
6. Prepare a draft literature survey addressing:
   a) results, recommendations, approaches and guidelines from studies that are immediately useable by the Navy in the upgrade of the HALIFAX class *(Complete)*;
   b) problems in the design, development, and implementation of military tactical or operational decision support that have hampered or precluded success of these systems ("lessons learned") and how these problems might be avoided *(Complete)*;
   c) Issues in naval team decision support that require further research and the implications of not addressing them *(Complete)*.
7. Review draft with Scientific Authority, revise, and deliver Final Report. *(Complete).*

1.4 Approach to the Literature Survey

1.4.1 Interpretation of Goals

The overall goal of the literature survey is to identify recommendations and guidelines that can be put to direct use in the upgrade of the HALIFAX class. The main thrusts of the survey are toward decision support and C2. Thus, the review considers decision support broadly as systems that enhance human performance in C2 rather than just in terms of computer systems that enhance information processing. This approach is consistent with ecological approaches to human cognition and system design (see Webb et al., 1993).

We believe that recommendations and guidelines can be taken from four broad areas:

- Theory.
- Methodologies.
- Empirical Results.
- Guidelines.

Theory. This survey will review relevant theory regarding decision making, situation awareness, team performance, training and expertise, and human-computer interaction. The purpose of this section is to provide an understanding of the issues surrounding decision support for naval C2 and identify the cognitive processes of human decision makers. The Theory section will identify the major needs for decision support and indicate, generally, how best to develop decision support.

Empirical Results. The section on Empirical Results will outline major findings pertinent to the theoretical issues raised in the first section. The goal is to review findings that indicate the applicability of the theories discussed in the first section to decision support for the HALIFAX class.

Methodologies. The section on Methodologies will discuss ways to further study issues, especially those concerning decision making, related cognitive processes, and the evaluation of decision support systems in the specific context of the HALIFAX class. Theory is valuable only if researchers have valid and reliable methods to resolve theoretical issues.

Guidelines. The final section will review a number of existing design guidelines for decision support and other related areas. These guidelines contain procedures and criteria for effective design of support for decision support, Situation Awareness (SA), Human-Computer Interaction (HCI), and team performance. They are, in a sense, compendiums of lessons learned. These guidelines are drawn from a number of sources, including civilian industry, which has invested greatly in HCI.

In all sections, our approach is to describe existing theories, methods, and results, compare them, and offer recommendations. Thus, the survey will serve as an overview of many different perspectives but also as an analysis each perspective to identify ideas, methods, and findings that can guide the development of modern C2 decision support. A set of specific recommendations is provided at the end of each section. These recommendations derive directly from the content of each section. Some are general and others specific but all recommendations indicate lessons learned, theoretical implications, or needs for further research. Because the review sampled diverse literature, many recommendations will need to be validated for the HALIFAX upgrade.

Overall, our approach is to provide both theoretical analysis and practical guidance. Each section will contain recommendations directly aimed at the HALIFAX class upgrade.

1.4.2 Summary Sections

The review covers a large literature, which reveals many issues and concepts. Furthermore, these issues are neither simple nor always well defined. Thus, the state of the literature itself can act as a barrier to understanding decision making in the naval context.

Another barrier to understanding is the overlapping nature of issues related to decision making and C2. We will discuss topics ranging from decision making to SA, from SA to teamwork, from teamwork to expertise, and so on. None of these topics, however, can really be understood in isolation. Decision making, for example, depends on SA; without an understanding of the situation, one cannot retrieve a relevant solution or define the problem space. SA, however, depends on decision making to set goals and identify relevant situational factors; without a goal and sense of what is important, one cannot develop complete SA. Consequently, it is important to consider how all these topics might relate to actual conditions and mission performance in naval operations. It is also important to note that our discussion is preliminary and speculative, intended as illustration only. The discussion may, however, suggest how to bridge the gap that sometimes exists between the perspectives of the operational user and the researcher.

To help readers, we include summary sections that review the issues raised in the context of general naval operations. Given the breadth of the literature review, we have avoided considering specific mission scenarios and instead examine a general, schematic outline. In this way, we can consider all issues and see how mission contingencies can affect decision making. This section will follow the broad outline of naval operations, establishing an appropriate context and stages within an operation. At each stage, we will discuss/illustrate:

- Relevant issues.
- Important themes and concepts.
- Major questions raised by the literature review.

The discussion will follow three broad stages of an operation:

- Mission planning, including planning and rehearsal.
- Coming on watch, including watch transfer.
- On watch surveillance, threat assessment, and threat response (implementation).

In addition, we will initially consider the general issue of the context of operations. The context affects all aspects of decision making so that it is impossible to describe or analyze decision making in detail without knowing the context in which decision makers operate.

1.4.3 Layout of the Literature Review

As mentioned above, this literature review is organized around four main subjects, theory, empirical results, methodologies, and guidelines. Each subject is discussed in its own section. In addition, this introductory section establishes the necessary background to understand and interpret the results of the literature review. A final section summarizes the major points of all sections and provides a final set of conclusions. Summary sections follow the major content areas. The first summary deals with both theory and empirical results. Given the close tie between theory and empirical results there seemed little value in summarizing each separately; in fact, considering the two together helped to illustrate major issues. Subsequently, there is a summary section for Methodology and for Guidelines.

Thus, the layout of this report is as follows:

1. Introduction.
2. Theory.
3. Empirical Results.
4. Summary of Theory and Empirical Results in the Operational Context.
5. Methodology.
6. Summary of Methodology in the Operational Context.
7. Guidelines.
8. Summary of Guidelines in the Operational Context.
9. Summary and Conclusions.

1.4.4 Context

Given the goal of supporting the HALIFAX class upgrade, this survey will examine the literature with respect to the immediate concerns of the Canadian Navy and, in particular, the role of the CPF. This is not to say that we will exclude material that does not directly refer to the CPF but, rather, that we will evaluate the pertinence of material in terms of its ability to identify issues, offer theoretical insight, describe potentially useful methods, discuss potentially relevant findings, and provide practical guidelines for designing decision support in the context of C2 for the HALIFAX class.

Thus, most of the literature reviewed in this report discusses decision support and related issues in the naval environment. Furthermore, we have sought literature that deals with particular aspects of the naval environment that are especially relevant to the role of the CPF. The primary aspects of this role can be summarized as:

- Operate as single ship and as part of a task group.
- Serve in multinational operations
- Engage in primarily defensive actions

- Deal with air, surface, and subsurface threats
- Operate in littoral settings
- Operate in joint (land, air, sea) taskforces.
- Operate with non-military agencies in non-military operations (humanitarian aid, peacekeeping, policing, etc.).
- Be flexible

The last aspect is probably the most crucial given the rapidly changing nature of the global political situation and the limited resources of the Canadian Navy. The CPF will probably be called upon to serve in unanticipated roles in the future. Thus, this review has considered the need for decision support in novel contexts.

A significant proportion of the literature reviewed comes from United States of America (US) research establishments and deals with US naval vessels, missions, and operations. As discussed below (Section 1.5.2), this partly reflects the state of the literature, which contains a very large number of US publications. There is, in addition, reason to consider US sponsored research heavily in the review. The United States is the major ally of Canada and the Canadian Navy is very likely to serve with the US in future operations, often in a subordinate role. Thus, Canadian forces will be called upon to coordinate with the C2 structures of other nations, especially that of the US.

1.4.5 Assumptions

To best achieve our goals, we believe that the literature review should take a broad view of theory and empirical results. That is, the review should attempt to identify useful theory and results, providing an overview of research that has been done in the areas relevant to the CPF. Thus, part of what the review will do is provide an unbiased account of competing theoretical perspective. The value of this is that it will frame specific questions and issues to be resolved in the HALIFAX class upgrade. It is important to understand *why* certain theories or findings provide guidance for designing decision support for the HALIFAX class. This can only be done by understanding the competing views in the literature.

This review will focus on the C2 functions of the Operations Room (OR) team aboard the CPF. This focus is consistent with the purpose of the review, to develop recommendations regarding single-ship C2. We will focus in particular on the "backrow" team, consisting of the Commanding Officer (CO), Operations Room Officer (ORO), Sensor Weapons Controller (SWC), and Assistant Sensor Weapons Controller (ASWC). These four positions are the focus of the OR team's decision making responsibility and are in greatest need of decision support (Webb & McLean, 1997). Other important roles include Operations Room Supervisor (ORS), Ship Air Controller (SAC), Officer of the Watch (OOW), and Combat Systems Engineer (CSE).

1.4.6 Definitions

This review will examine research from a number of countries that addresses decision making at all levels and all positions in the OR. It will also consider issues related to decision making, such as SA, team performance, and interface design. Thus, the review will consider members of the OR in many different ways. It will consider them as decision makers who are attempting to determine a Course of Action (COA), as members of a team, as users of

particular pieces of equipment, as operators of computers and other equipment, and as performers of specific tasks and actions. Consequently, different terms will be used throughout this review to refer to the same individuals – members of the OR team. The multiple terms, however, are necessary to preserve the different aspects of OR team function and C2 that are under investigation. Table 1.1 below summarizes the terminology used to refer to individuals in this report.

Term	Definition	Who This Term Refers to
C2-Related Terms		
Decision maker	A person engaged in mental activity to determine a COA, identify a contact, or make a recommendation	Individuals at any position in the OR
Commander	A person who exercises command at some level	Individuals with command authority in the OR
Individual/Team-Related Terms		
Operator	A person who is operating some piece of equipment or system	Individuals at any position in the OR
User	A person who currently works with, or will work with in the future, some piece of equipment or computer system	Individuals at any position in the OR
Member	A person who belongs to a team	Individuals at any position in the OR
Performer	A person who performs some task or action	Individuals at any position in the OR
Participant	A person who is serving in an experiment or study	Individuals selected for an experiment or study

Table 1.1. Terminology used in this report.

The terms "Commander" and command" are, in particular, used ambiguously in the literature. Typically, the term commander is used to denote the person in authority, responsible for decisions and setting the framework in which decisions are made. Usually, this term is reserved for a certain high level of aggregation, such as a naval ship, group of ships, or group of aircraft, or all forces in a certain area.

The term commander is used to refer to the CO aboard the CPF. Unfortunately, it also sometimes appears as a term referring to other positions aboard US Navy vessels. For example, Canadian Task Groups can have Anti-Air Warfare Commander (AAWC) and US Navy vessels can have a Composite Warfare Commander (CWC). In this case, commander refers to a position in the OR and not the rank or command responsibilities of the individual. In addition, individual command functions may be delegated or assumed in whole or in part by several OR positions. This highlights that C2 is not in the hands of a single person but is distributed across a team, which makes it inaccurate to talk of *the* commander when discussing the whole of C2 functions.

The term commander is also sometimes used in the literature to simply refer to an individual making decisions and issuing directives in a scenario or experiment, regardless of the true rank or position of the individual. Rather than disentangle these different uses of the term commander, we have adopted the term commander (with the lower case spelling) as a general term referring to an individual who issues directives at some level in a command structure. A

commander need not be the CO because discussion can focus on directives issued at lower levels of command. The term CO has been reserved strictly for the commanding officer of a Navy vessel or task group.

Unfortunately, the term commander frequently implies that one person undertakes all decision making activities. This is not the case as the CO or senior authority typically accepts or rejects the COAs recommended by subordinates. These recommendations themselves are based on decision making activities of the subordinates and OR team. Thus, command is involves distributed decision making activities and a shared mental model of OR functions and mission objectives. This shared mental model is a set of assumptions (implicit or explicit; see Pigeau & McCann, 1998) established in the context of common training and planning and preparation (rehearsals, briefings, etc.) conducted jointly by team members in the days or weeks leading up to mission execution. The role of the CO is to employ resources (equipment, software, and personnel) in terms of both generic and mission-specific capabilities to ensure that the right decisions are made. The role of the CO is not to perform all decision making activities themselves.

Consequently, the terms commander and command are loaded with ambiguity, yielding several different interpretations. This review takes the general focus that can be termed distributed command and control. Decisions are viewed as distributed across the OR team and over time, particularly over the period of planning and preparation. The nature of the distribution of decision making activities will be crucial to determining the re-allocation of C2 functions between personnel and technologies in the HALIFAX class upgrade. It is important that the analysis of decision making, and hence decision support needs, is not confused by ambiguities inherent in terminology. It is important that decision support not be narrowly focussed, explicitly or implicitly, on the commanding officer or within the OR Command Team.

1.5 Limitations

1.5.1 Extent of the Literature Review

A vast literature exists on decision making, tactical C2, and other topics relevant to this report. To ensure that the literature review remained a manageable project, limits were placed on the literature to be surveyed. The search process is described in detail by Bryant, et al. (1998), who developed a set of keywords to identify relevant articles. The keywords fell into two categories, those pertaining to the naval environment and those pertaining to C2, information processing, and decision making issues. These keywords were used to search two databases, one provided by PsycInfo, which is operated by the American Psychological Association (APA), and the other by the NTIS.

The search procedure focused on identifying articles in the naval domain. Additional articles were identified through other means (e.g., through citation in another article, from DCIEM, in-house resources) and some pertain to other domains, such as army, air force, coast guard, and behavioural sciences. Nevertheless, the majority of literature surveyed directly addresses naval issues (although some address issues overlapping multiple domains). Further review should consider non-naval C2 and decision support issues.

We also restricted the age of articles considered in the literature review. In general, only articles published within the last 10 years were considered relevant. Furthermore, articles

published within the last 5 years were given priority. This rule was not applied inflexibly; some older articles were deemed significant despite their age and others were "classics." The restriction on the age of articles, however, was necessary due to the vast number of papers published over the years.

In addition to restricting the kinds of articles searched for, we restricted the number of articles considered. Approximately 200 articles were obtained for detailed review. These articles form the basis for this review. A number of other articles are referred to briefly.

1.5.2 Influence of American Research

One goal of the literature review was to survey research from Canadian, American, European, and Australian sources. Although this goal was achieved to some extent, the overwhelming majority of articles obtained were of American origin and researched issues in the context of the US military. This reflects, in part, the fact that the United States possesses greater scientific and military resources than other nations and has produced more literature in this area. US research was a valuable resource for this review and provided the most extensive coverage of issues surrounding naval C2 and decision support. In addition, the US Navy is concerned with many of the same aspects of the naval environment as the HALIFAX upgrade, namely operations dealing with air, surface, and subsurface threats in littoral settings. The heavy weighting of US research also reflects, in part, the databases we searched. Both the PsycInfo and NTIS databases contain articles from numerous countries but their focus seems to be on US publications.

1.5.3 Data Fusion

A major concern for the upgrade of the HALIFAX class will be providing an appropriately unified tactical picture to OR team members. Thus, the topic of *data fusion* will be very important. The ability to combine information from numerous sensors has the potential to dramatically enhance individual and team Situation Awareness (SA), data processing, and decision making. Currently, most research in this area deals with technological issues and few articles are exclusively devoted to Human Factors (HF) considerations. As a result, data fusion was not discussed as a separate topic in this report. Instead, it was dealt with in relation to other topics. Discussions of decision making, SA, and so on highlighted the need to synthesize multiple data sources and provide an enhanced picture of the tactical situation.

1.6 Acronyms and Abbreviations

The following acronyms and abbreviations are used in this report:

2D	Two Dimensional	KSA	Knowledge, Skill, Attitude
3D	Three Dimensional	LTM	Long-Term Memory
AAW	Anti-Air Warfare	MAUT	Multi-Attribute Utility Theory
AAWC	Anti-Air Warfare Coordinator	MCA	Multi-Criteria Analysis
AESOP	Aircrew Evaluation Sustained Operations Performance	MDS	Multi-Dimensional Scaling

AI	Artificial Intelligence	MOE	Measure of Effectiveness
APA	American Psychological Association	MOFE	Measure of Force Effectiveness
ASWC	Assistant Sensor Weapons Controller	MOOTW	Military Operation Other Than War
BATMAN & ROBIN	Battle Management Assessment System and Raid Originator Bogie Ingress	MOP	Measure of Performance
C&D	Command and Decision	MSC	Mission Success Criteria
C2	Command and Control	NDM	Natural Decision Making
C3I	Command, Control, Communication, and Intelligence	NTIS	National Technical Information Service
C4I	Command, Control, Communications, Computers, and Intelligence	OODA	Orient, Observe, Decide, Act
CTBF	Critical Team Behaviours For	OR	Operations Room
CDM	Critical Decision Method	ORO	Operations Room Officer
CDS	Command Decision Support	PFC	Point For Consideration
CGA	Conceptual Graph Analysis	PMD	Plan, Monitor, Direct
CIC	Combat Information Centre	R&D	Research and Development
CIM	Critical Incident Method	RC	Required Capabilities
CM	Conceptual Mapping	ROE	Rules of Engagement
CO	Commanding Officer	RPD	Recognition Primed Decision
COA	Course of Action	SABER	Situation Assessment By Explanation-based Reasoning
COGNET	Cognition as a Network of Tasks	SAC	Ship Air Controller
CT	Command Team	SAGAT	Situation Awareness Global Assessment Technique
CPF	Canadian Patrol Frigate	SHOR	Stimulus, Hypothesis, Option, Response
CTA	Cognitive Task Analysis	SME	Subject Matter Expert

CTT	Critical Thinking Training	STEP	Story, Testing, Evaluation, Plans
CWA	Cognitive Work Analysis	SWC	Sensor Weapons Controller
CWC	Composite Warfare Commander	TADMUS	Tactical Decision Making Under Stress
DCIEM	Defence and Civil Institute of Environmental Medicine	TANDEM	Tactical Naval Decision Making System
DEFTT	Decision Making Evaluation Facility for Tactical Teams	TAO	Tactical Action Officer
DR	Decision Requirement	TARGETS	Targeted Acceptable Responses to Generated Events or Tasks
DSS	Decision Support System	TDM	Tactical Decision Making
ES	Expert System	TDS	Tactical Display Station
EWS	Electronic Warfare Supervisor	TIC	Tactical Information Coordinator
FSA	Finite State Automata	$TIDE^2$	Team Interactive Decision Exercise for Teams Incorporating Distributed Expertise
GT-AAWC	Georgia Tech Anti-Air Warfare Coordinator Simulation Suite	TMT	Team Model Trainer
GUI	Graphical User Interface	TODAINFO	Total Defence and Information System
HCI	Human-Computer Interaction	TPAB	Team Performance Assessment Battery
HEAT	Headquarters Effectiveness Analysis Tool	TPM	Task Process Model
HF	Human Factors	US	United States (of America)
HMD	Head Mounted Display	VCC	Virtual Command Centre
IDS	Identification Supervisor	WG	Working Group
IFM	Information Flow Model	WM	Working Memory
IW	Information Warfare	WWW	World-Wide Web

2. Theory

2.1 The Naval Tactical Environment

Virtually every theoretical perspective on SA, team performance, and decision making indicates the importance of the environment and task (and hence the distributed nature of cognition in team settings) to understanding human performance. The Naval tactical environment poses special problems for the human decision maker as well certain constraints and task demands. Thus, understanding the nature of the task environment is essential to evaluate theories of individual and team performance. Unfortunately, the Naval tactical environment is not stable and has been changing in recent years.

2.1.1 Command and Control

This survey focuses on Command and Control (C2) functions. C2 is the exercise of authority or direction by a commander over assigned forces to accomplish a mission (JCS, cited in Stevens et al., 1996). As such, C2 encompasses a wide range of specific tasks. Most of these tasks centre upon decision making and the use of information. A related concept is that of Information Warfare (IW), which consists of actions taken to exploit, manipulate, or destroy an enemy's information resources while defending one's own information and information systems (Stevens, et al., 1996). The main principles of C2 according to Sundin (1996) are to:

- "Get inside" and understand your opponent's decision making process.
- Accurately assess and understand the tactical situation.
- Identify and classify friendly and enemy units.
- Predict and anticipate the actions of the enemy.
- Initiate the proper actions at the proper times based on available information.

The commander and officers of the Operations Room (OR) of the Canadian Patrol Frigate (CPF) are very much highly skilled information processors. Their success in fulfilling specific missions will depend on how well they take in and process information and how well they use that information to decide on a Course of Action (COA).

2.1.2 Issues in Defining C2

The description of C2 above is a fairly standard one and serves as an orientation to the survey of the broad literature. The concept of C2, however, is open to different interpretations. Pigeau and McCann (1995) have argued that typical definitions of C2 fail to distinguish two key concepts. They note that, under the NATO definition, *command* is the authority vested in an individual to direct and control military forces. It is associated with attributes of the individual, such as responsibility and leadership. *Control*, on the other hand, is the authority exercised by a commander over the activities of subordinate organizations. It is associated with processes and functions of systems, such as communication. Broadly stated, command can be thought of as the authority to issue commands and control as the means by which those commands can be enacted.

Putting the concepts of command and control together in C2 has led to an emphasis on control and technology, and the neglect of command (Pigeau & McCann, 1995). Researchers often

treat the human as a user of a C2 system rather than as a directive agent. Pigeau and McCann (1995) argue that this distorts the true picture of tactical decision making. In their view, command is the more important concept. Issues of command, such as responsibility, authority, motivation, and teamwork, must be addressed before issues of control can be considered. Thus, C2 is exercised by an individual, "the commander", rather than a team (this is true also in commercial and military aircraft where the pilot exercises authority).

In this view, C2 can be redefined as *"The establishment of common intent and the transformation of common intent into co-ordinated action"* (Pigeau & McCann, 1998). Under this definition, the focus is on the human commander and on human teams. In particular, important considerations are the explicit and implicit transmission of intent from a commander to a team. For Pigeau and McCann, effective C2 stems as much from shared implicit and explicit intent among the Command Team (CT) and within the unit/ship as a whole as on the existence of hardware, software, and automated algorithms. Furthermore, there will be an increasing dependence on individual "experts" to act independently but in consort to arrive at coordinated team decisions that fit an overall strategic intent for the force as a whole due to rising educational levels, arcane technological issues underlying COA choices, and short time lines. In a previous study (Webb & McLean, 1997), interviews with CPF command teams repeatedly emphasized the difficulty of achieving comprehension across specialist dividing lines (e.g., SWC versus SAC) and increasing dependence on specialist advice, depending of an Operations Room Operator's (ORO's) or Commanding Officer's (CO's) specialization.

In contrast to Pigeau and McCann's view, many researchers, including those in the United States of America (US), have taken a different perspective on C2. Rather than trying to distinguish concepts within C2, they have sought to broaden the scope of this concept. It is common now to see researchers refer to C4I (command, control, communications, computers, and intelligence) rather than C2. The purpose of this redefinition seems to be to capture the tight inter-relation of all aspects of tactical command. Because issues of command and control depend on communications, computers (and other equipment), and intelligence, researchers have sought to analyze the entire tactical situation as a system.

There seems to be value in both these approaches to C2. Certainly, the emphasis on humans in command teams is more warranted than a focus on technology. Technology will have value only with respect to the goals and needs of the members of command teams. However, it is also important to keep sight of how the human commander will have to work within a complex system that contains not only control processes but also communications and other processes.

2.1.3 Perspectives on C2

There are a number of perspectives or schools of thought on C2. These include the Plan, Monitor, Direct (PMD) model (Canadian Army) and the Orient, Observe, Decide, Act (OODA) model (Canadian Navy). Although these perspectives differ in some respects, we have purposely avoided taking a position on which particular model might be best. There are a number of reasons why adopting a specific model is inappropriate for this literature review. First, the purpose of the literature review is to explore a range of approaches and theories related to C2 and decision support. Adopting one model would limit or bias the search for relevant literature. Second, we wish to take a broad view of C2 that reveals major issues in C2 and relations between research in C2, decision making, and other pertinent areas. Finally,

despite differences in approaches, perspectives on C2 often share many important features. PMD and OODA, for example, share a general focus on acquiring information and directing action based on a good understanding of the situation. Thus, there seems to be more to gained in the literature review by examining C2 at a general conceptual level rather than accepting one particular view.

2.1.4 C2 Theory and Practice

C2 can be exercised at three levels, strategic, tactical, and procedural (Flin, 1998).

Strategic. This level is concerned with overall policy. It focuses on long-term initiatives and the distribution of resources. The goal is to set priorities and Rules of Engagement (ROEs) for implementation at lower levels of command. Projection and prediction are crucial. There is a strong emphasis on analytic procedures (see Section 2.2.3.1) and the use of formal models when possible. There is a greater need here than at lower levels to achieve an optimal solution to problems because decisions made at this level will have more far-reaching implications. Typically, strategic considerations are determined in advance of a mission, under much less time pressure, and with resources to gather and process information across a broad range of topics. The strategic level may be seen as establishing a framework of intent in which tactical goals are set. This implies that a CT must keep the strategic model or picture in mind as part of SA and decision making as tactical actions are taken.

Tactical. This level is concerned with planning and coordinating actions determined at the strategic level. Tactical command is goal-oriented but focuses on following established procedures to accomplish the mission at hand. SA is important to ensure procedures are implemented only when appropriate. In addition, a commander or CT needs to be able to adapt strategic elements to fit the particulars of the situation. Typically, commanders face severe time pressure, which forces them to concentrate on obtaining working solutions rather than optimal solutions. Hence, there is greater reliance on intuitive rather than analytic decision making processes.

Procedural. This level is actually referred to by Flin (1998) as the operational level. It implements orders from the tactical level. It is rule-based and, like the tactical level, depends on SA and intuitive (recognitional or experientially-driven) decision making. Time pressure is typically very great. For these reasons, we prefer the term "procedural" to avoid potential confusion with theatre-wide operations

This review will focus on the tactical and procedural levels of C2 because they are more relevant to the members of the OR team of the CPF. In addition, changes in the nature of the naval environment will place greater responsibilities and burdens on tactical and operational command (e.g., Hughes, 1996; O'Neill, 1996). The strategic level cannot be completely neglected, however, because it pertains to pre-mission planning, which will establish principles, knowledge-bases, and ROEs for the tactical and procedural levels. It is, in effect, an implied mental framework from which all other C2 decisions will flow and against which they must be judged for all members of the CT. If members of the OR CT have differing mental models of the strategic intent, effective C2 at the tactical and procedural levels will become less likely. Furthermore, the strategic and procedural levels will be subject to change (e.g., if own forces are attacked then the strategic aim may shift quickly from reactive

defence to pro-active and pre-emptive destruction of enemy forces). The point and degree of change may not be easy to anticipate or plan.

Within the tactical and procedural levels of C2, there are at least four management tasks that must be performed (Anderson, 1990). These tasks represent major functions underlying C2. Each of these tasks is an issue for decision making and support and all contribute to improved information management in a cognitive sense. The tasks are:

- Information management (e.g., sensor data).
- Platform management (e.g., ship maneuver).
- Resource management (e.g., fuel, weapons).
- Tactical management (e.g., situation assessment).

Information management. This is made difficult by the tremendous volume of data that is presented to the commander. At any given time, a ship may be tracking dozens or even hundreds of air and seacraft, especially in a littoral context. Another issue is the ambiguity of data. Information can be uncertain or open to multiple interpretations. The primary Human Factors (HF) concern of information management is to reduce the volume and ambiguity of information to a level compatible with the capabilities of the human operator while maintaining oversight of all craft within sensor range.

Platform management. This requires an integration of sensor data and data on ship position and status. Further, the tactical team must interact with the bridge team to exert effective control of the ship. Primary concerns here are developing a unified picture of the tactical situation, including enemy, friendly, and neutral craft, one's own ship, aircraft, and geographical features. In addition, the effectiveness of communications is an issue.

Resource management. This requires integration of own-ship and track data as well as monitoring of ship status. The effective management of ship resources contributes to all other areas of management.

Tactical management. This is an overall control function subject to all of the above concerns. It encompasses concerns of mission effectiveness and response readiness. Thus, the command team must gather information, identify friendly, neutral, and hostile forces in the area, determine how to achieve mission goals, and respond appropriately to any threats and unexpected events. An overall concern, then, is to create a comprehensive and accurate understanding or picture of the local situation and relate that understanding to mission requirements.

These broad tasks are manifested in a large number of specific tasks. For example, the US Naval Publication 6.0 (cited in Sovereign, 1996) indicates that mandatory C4I capabilities include collection, manipulation, and transmission of data, providing warnings and attack assessments, and tracking and control of assets (see also Mason, 1995).

2.1.5 Littoral Warfare

Increasingly, the changing global political situation is forcing changes in Naval roles and operations. In the emerging global situation, there are no clear military threats. Instead, threats will emerge from non-traditional sources, such as regional conflicts (O'Neill, 1996). Increasingly, nations are adopting "total defence" strategies involving increased cooperation among land, air, and sea forces, and between international forces, inside and outside NATO

(Sundin, 1996). In addition, military forces are being called upon to serve in cooperation with civilian agencies (e.g., humanitarian relief efforts) (O'Neill, 1996).

One implication of these changes in the Naval Warfare environment is that Canadian Naval forces can expect to operate increasingly in littoral (near land) situations. Littoral warfare, however, poses a number of problems for effective C2 not faced in open-water conflicts (Hutchins, 1997). These include:

- Presence of friendly, neutral, and hostile shipping and aircraft.
- Congested water and air spaces.
- Complex ROE.
- Increased potential for collateral damage to friendly, neutral, or civilian resources.
- Uncertain motivations of forces in the area.
- Speed of development of events.
- Operate in multi-vessel groups.
- Participate in subordinate roles in multinational operations.
- Participate in multi-lingual groups.
- Participate in operations with multiple goals.

The littoral environment has two important implications for naval operations (Hughes, 1996). First, there is necessarily close interaction of air, land, and sea forces. The naval CT must be prepared to work with friendly land and air units but also to defend against enemy land and air forces. Second, the congested nature of littoral waters increases the difficulty of most C2 functions. Congested waters, especially in industrial or quasi-industrial regions (presence of fishing, oil platforms, etc.), increase the opportunities for enemy concealment while making it more difficult to detect, track, and identify surface and subsurface vessels and aircraft in the area. Meanwhile, one's own forces can be more effectively monitored by the enemy. The lack of battlespace can seriously constrain maneuvers. The presence of neutral and civilian craft increases the complexity of ROEs and requires great care to avoid civilian and friendly casualties. In short, CTs will face greater challenges in future missions.

2.1.6 Missions and Doctrine

The shift to littoral warfare has implications for the kinds of missions CTs will undertake in the future. The main focus of littoral warfare is (Stevens et al., 1996):

- Creating secure areas ashore.
- Moving troops and assets ashore.
- Withdrawing troops and assets.
- Land warfare.
- Creating a presence in protecting sea-borne assets.

Commanders of CPFs can expect to play a role in these kinds of operations as well as patrol and escort missions. Within these classes of missions, the 3 primary tasks of sea control will be (Foster, 1992):

- Anti-air warfare (neutralizing enemy air platforms).
- Anti-submarine warfare (neutralizing enemy subsurface platforms).
- Anti-surface warfare (neutralizing enemy surface platforms).

Consequently, much of the OR team's activities will be aimed at detecting and responding to air, surface, and subsurface threats. Threats, of course, can switch between these missions (e.g., as a ship fires an anti-ship missile). As noted above, the littoral environment makes these tasks especially difficult.

Despite the increasing importance of littoral operations, and the concomitant increase risk to Naval forces, Naval doctrine is not well developed in this area (Sovereign, 1996). NATO forces developed doctrine in the context of a perceived Soviet/Warsaw Pact threat of large-scale, deep water conflict and there is a need now for new tactical capabilities, an understanding of new opposition tactics, and new training programs (US Tactical Training Manual, Ch. 2). As it currently stands, there is currently a shortage of practical guidance to help commanders act in ambiguous situations with the speed necessitated by littoral environments.

2.1.7 Implications for Decision Support

Hughes (1996) notes that previous US doctrine has focused on the strategic and operational levels. He advocates greater concentration on the tactical level, which will provide greater benefits in the littoral environment. The US Tactical Training Manual (Ch 2.) has established some general needs in this area. Specifically, commanders should:

- Be able to act quickly, based on minimal planning and preparation.
- Respond within the ROEs.
- Have fast and accurate communications.
- Be able to monitor all aspects of the theater, including air and land.
- Be able to focus on multiple threats.
- Be able to coordinate with other platforms.
- Be able to operate in unfamiliar and ambiguous situations.

The last of these requirements poses the biggest challenge; commanders will be expected to act in situations beyond their experience. The fundamental task of tactical C2 can be considered to be reasoning with uncertainty in time-constrained situations (Hair & Pickslay, 1992). Thus, commanders will have to rely increasingly on decision support. In this review, decision support will most often refer to computerized aids to support decision making. The other requirements also point to decision support needs. Aids will be needed to increase the speed of decision making, provide quick and constant access to information such as ROEs, speed and improve communications, expand the tactical picture, enhance commander's ability to form SA, and help the commander simultaneously process multiple data streams.

Satisfying all of these needs will be a tall order. An understanding of human cognition and the C2 domain, however, can advance the cause. For example, the C2 requirement to monitor all land, sea, and air traffic in an area point to the criticality of SA to performing C2 functions. Humans, however, have limited attentional, memorial, and cognitive resources to apply to the task. Thus, the heavy workload associated with C2 necessitates some form of support that reduces the effective demands of maintaining accurate perception of the situation, identifying problems and threats, recognizing when there is a need for action, and planning actions (Hutchins, 1996).

Stevens et al. (1996) offer a list of seven key areas need to be addressed to provide a good model of C2. This list indicates what topics should serve as the focus for research on specific means to enhance C2 effectiveness. The major aspects of C2 are:

- Collection of information, uncertainty levels, equipment status, and risk assessments.
- Dissemination of information.
- Data fusion to facilitate extraction and correlation and the formation of a unified tactical picture.
- Operator displays (the form in which fused data are presented).
- Situation assessment.
- Decision making.
- Metrics to quantify the above aspects.

2.1.8 Recommendations

- Prepare to operate within complex settings, including single ship and task group, uni-national and multinational, and littoral and deep-water settings.
- View C2 as team activity rather than the role of the commander as an individual.
- Tactical and procedural levels of C2 should remain the focus of R&D for the CPF but some effort should be devoted too the strategic level.

2.2 Decision Making

Making decisions is perhaps the most important task of all members of the OR team. The commander, in particular, must decide how to respond to potential threats and determine a COA. This section examines the major theories of decision making as applied to Naval C2 and other pertinent factors that affect decision making performance.

2.2.1 Decision Making and Problem Solving

Technically, there is a distinction between *decision making* and *problem solving*. Decision making refers to the selection of an option or COA from a multiple set of well-defined alternatives (e.g., Sternberg, 1996; von Neumann & Morgenstern, 1944). Problem solving is the achievement of some goal state given some initial state, where the computational path between initial and goal states is not obvious (Mayer, 1989). In practical terms, however, this distinction is not particularly useful. First, tactical decision making can be viewed as a form of problem solving. The decision maker is in an initial state in which he or she is confronted with a threat or potential threat and must determine how to neutralize it. This description is especially apt for the tactical situation because the alternatives for responding are rarely clearly specified and the decision maker must compute them as part of determining a COA. Second, although a tactical decision maker rarely has well-defined alternatives to choose from, he or she does have a limited number of possible COAs. Furthermore, it is the selection of a COA that is the critical outcome of the tactical decision; other decisions such as identifying a target only serve to help the decision maker select a COA. Finally, studies of problem solving can illuminate the cognitive processes of tactical decision making. In particular, there is great correspondence between the formal definition of problem solving and the overall tactical task of gathering information, identifying threats and opportunities, and selecting a COA.

Gilhooly (1989), for example, has indicated three broad steps to solving problems:

- Detect that a problem exists (i.e. notice the discrepancy between one's current and goal states).
- Formulate the problem completely (i.e. define current and goal states and the rules for translating between them).
- Attempt to solve the problem (i.e. formulate a series of actions to achieve the goal state).

Considering these steps, we see that problem solving involves 1) taking in information to represent the situation and identifying threats or issues that must be resolved, 2) representing the situation or maintaining SA, and 3) identifying a response or COA that will neutralize the issue. Because the tactical crew operate in an environment where they are responsible not just for deciding on a COA but assessing the situation, classifying the problem, and then determining what to do, the tactical task is more akin to the classical definition of problem solving than decision making. Thus, in this report, we will view decision making as a problem solving activity, taking a very broad and high-level view of the cognitive processes involved. Greater insight into C2 can certainly be gained by examining the broader literature on problem solving.

2.2.2 Classification of Problems

There are several classifications of problems, summarized by Gilhooley (1989). First problems can be *well-defined* or *ill-defined*. Well-defined problems are ones in which all problem elements (objects, relationships, constraints, rules, etc.) are clearly specified. Ill-defined problems are ones in which these elements are not all clearly specified. Second, problems can be *adversarial* or *non-adversarial*. Adversarial problems are ones in which a person faces a rational opponent who attempts to undo his or her efforts, adding another layer of complexity. Non-adversarial problems are ones in which the person faces non-responsive materials. Third, problems can be *systematically rich* or *systematically impoverished*. Systematically rich problems are ones in which the problem solver brings a great deal of knowledge to the task in the form of memories, experiences, and analogies. Systematically impoverished problems are ones in which the problem solver has little experience or knowledge to use.

In addition to Gilhooly's classification, we can also distinguish between *individual* and *team* problem solving. Theories of team problem solving or decision making will be discussed in detail later (see Section 2.5), but team problem solving entails coordinating one's efforts with those of others, adding additional cognitive burdens (although enabling opportunities).

Problems can also be *short-term* or *long-term*. Short-term problems must be dealt with in short periods of time either because the problem is straightforward or because there is limited time available to solve the problem. Long-term problems allow ample time to solve the problem and are generally complex.

By this classification scheme, tactical problems are typically ill-defined, adversarial, systematically impoverished, and team-oriented. As such, tactical problems are among the most difficult kinds of problems. They place the greatest number of constraints on the problem solver while providing the fewest resources or enablers for the problem solver to work with. In particular, people faced with tactical problems will need support in gathering information, representing the problem, and coordinating efforts with other team members and

also need a database to supplement their own lack of knowledge with the particular problem faced. Tactical problems are often short-term, demanding rapid response. Not all tactical problems, however, should be classified as short-term. Achieving mission goals may be a complex process requiring extended effort.

2.2.3 Major Distinctions in Theories of Decision Making

The major distinction in the study of decision making is between *analytic* and *intuitive* theories. There are numerous specific models within each broad class but they share certain core principles that set one class apart from the other.

2.2.3.1 Analytic Theories

Analytic theories have the longer history. They arise from the view of human cognition that described humans as information processors and active planners (Rouse, 1980, pp. 119-135). The emphasis in explaining human decision making and problem solving is on identifying how people take in information, code it symbolically, manipulate symbolic representations, and generate some output. As planners, people deal with problems by generating alternative plans, predicting consequences, and choosing one COA based on a rational analysis of the value of each COA.

Thus, a core principle of analytic theories is that the goal of decision making is to reach an optimal decision. Optimality is a difficult concept to operationalize but it is generally defined in terms of maximizing benefits such as enemy units destroyed and friendly units preserved in tactical situations. A second principle is that decision making involves an analysis of all available data and the evaluation of all possible hypothesis (Hutchins, 1996). As a consequence, decision making entails extensive computations and, hence, a great deal of time. Generally, analytic decision making involves the following basic steps (e.g., Hutchins, 1996):

- Specify all relevant features of the problem.
- Identify the full range of options.
- Identify key evaluation dimensions.
- Identify weights for each dimension.
- Rate each option on each dimension.
- Tabulate the results.
- Select the best option.

Key to this approach is the notion of a formal comparison (Klein, 1997). Analytic theories rely on a deliberate and procedural analysis to quantify alternative COAs. This assumes that all pertinent factors can be a) identified, and b) quantified in terms of their absolute or relative impact. Methods of comparison are based on formal logic and/or mathematical algorithms (Best, 1995), so that there is a direct link between facts and conclusions. The advantage of this is that the decision making rules are verifiable and valid. Also, the procedure for reaching a decision can be readily explained and documented.

There are numerous specific procedures for comparing alternatives. Many, for example, are based on Bayesian statistics and evaluate options in terms of base rates for different hypotheses and probabilities of accuracy of different observations (Klein, 1992). Zsambok (1992, cited in Klein, 1992) identified 15 specific analytic strategies, including

subjective expected utility, single feature difference, and elimination by aspects. With subjective expected utility, one computes the maximum utility, or benefit, possible based on subjective utilities or values and subjective probabilities assigned to events. With elimination by aspects, one focuses on one attribute of the alternatives at a time (Tversky, 1972a, 1972b, cited in Sternberg, 1996). One then forms a minimum criterion for that attribute and eliminates all alternatives not meeting the criterion. Another attribute is selected for the remaining options and the process repeated until only one option is left. Another analytic strategy identified in the context of tactical decision making is the use of successive pairs (Klein, 1992). Here, the decision maker performs successive pairwise comparisons of alternative COAs, choosing the better option each time.

In addition to rational comparison, other important concepts are that of *problem space* and *search*. These concepts come from studies of problem solving but have been applied to the problem of selecting a COA. The problem space consists of the various states of the problem (Anderson, 1995). A state is a representation of the arrangement of problem elements - the objects, conditions, and so on that make up the situation. The initial state is the representation of the problem before any actions have been taken, whereas the goal state is the representation of the desired configuration of problem elements. In this case, problem solving or decision making is viewed as a process of search - identifying the sequence of actions, or transformations, that will transform the initial state to the goal state through clearly specified intermediary states. These transformations can consist of any specifiable operation that will affect the problem elements in some way.

In analytic theories, problem space search is carried out by production systems (Anderson, 1995). *Productions* are rules for specifying and selecting transformations. They consist of a *goal*, some *application tests*, and an *action* that is performed if the conditions are met. Productions are very much like "if-then" rules. To solve a problem, productions are strung together to transform the initial state through the problem space to the goal state. Essentially, the problem solver must set up many subgoals corresponding to intermediary states, and determine the relevant productions (Best, 1995, pp. 436-442).

Two common ways of selecting the productions or operators are *difference reduction* and *mean-ends analysis* (Anderson, 1995). With difference reduction, one simply selects an operator that will reduce the distance in the problem space between the current and goal state. This has been likened to hill-climbing in which the problem solver chooses the operation that moves him or her down a gradient toward the goal state, given some constraint on how far one can move in any one step. This method can be ineffective if there are barriers, or *local minima*, between the current and goal states. The problem solver may reach a state in which there is no direct step toward the goal and must, in fact, move to a state further from the solution in order to ultimately reach the goal.

Means-ends analysis avoids this problem by beginning with an overall analysis of the problem space. The problem solver identifies a chain of subgoals between the initial and goal states. This sequence need not be a direct path. Rather, the problem solver focuses on the functions of operations and analyzes which functions are necessary to transform the current to goal state. Then the problem solver assigns concrete actions that will accomplish the functions.

Analytic theories of decision making assume that people perform rationally and with perfect computation (Sternberg, 1996, pp. 387-389). More specifically, they assume that decision makers are fully informed about all possible options and their attributes, infinitely sensitive to subtle distinctions among options, and always choose options that maximize utility. As a result, in studying analytic methods, researchers have focused on situations in which decision makers have sufficient time to generate options, evaluate each option, and select an action. Further, the consequences of errors or delay in decision making have not been immediately clear. Such situations bear little similarity to the fast paced, high risk environment of naval C2 raising questions regarding the applicability analytic theories to this domain (Klein, 1992, 1997; Hutchins, 1996).

2.2.3.2 *Intuitive Theories*

Although younger than analytic approaches, intuitive theories of decision making have become influential in naval C2. The rationale for applying intuitive theories is that analytic theories do not necessarily apply in natural settings where time is limited and data ambiguous (Klein, 1997). In particular, intuitive theories dispense with the concept of rational choice and assume that human decision makers use much less formal but much faster strategies.

Intuitive theories, or Natural Decision Making (NDM) theories as they are sometimes called, are based not on formal analyses but strategies that experienced decision makers, or so-called experts, actually use. Thus, it is an observational/descriptive approach that assumes human strategies are good strategies.

There are three basic principles underlying intuitive theories. The first is that decisions are made by *holistic evaluation* of potential COAs rather than by feature-by-feature comparison of alternatives. Typically, the decision maker does not even compare multiple options but only one COA at a time. Then a quick evaluation of the consequences and value of the COA is made. The evaluation process relies heavily on memory as the COA is compared to previous experiences and the outcomes recalled.

The second principle is that the decision maker relies on *memory/recognition* of COAs rather than an exhaustive generation and comparison of alternatives. It is assumed that decision makers must use their memory to generate potential COAs and the data used to evaluate them. A key concept is that of SA. The decision maker identifies potential COAs by first assessing the situation then recognizing past situations that are similar. From this experience, the decision maker can recall COAs taken in the past. The decision maker must also use memory of the outcomes of the previous experiences to determine the acceptability of potential COAs.

The third principle is that decision makers adopt a *satisficing* criterion rather than search for an optimal solution. Real world situations often demand very rapid responses and decision makers may have to accept a solution that merely works and not consider whether a better solution exists. It is assumed that decision makers stop their decision making process as soon as they identify an acceptable COA.

2.2.3.3 Comparison of Analytic and Intuitive Models

In comparing analytic and intuitive theories, we must consider both the intrinsic strengths of each as theories of decision making and their applicability to the naval C2 domain.

Intuitive theories have the advantages of being closely linked to what expert decision makers actually do in real-world domains and being applicable to dynamic, uncertain, and high risk environments. Their usefulness as models of decision making, however, can be limited by their informal nature. Because intuitive theories refer to memory and analogical processes, it can be hard to describe detailed steps or procedures. They offer a much less explicable account of the cognitive processes of decision making than do analytic theories. Also, intuitive theories are, by their nature, more descriptive than prescriptive. Although researchers often assume that expert decision makers engage in effective strategies (e.g., Klein, 1997), it is not clear that their performance could not be dramatically improved.

As mentioned, naval C2 is a domain requiring fast decisions based on limited and/or uncertain data, where there is high risk associated with potential COAs. Noakes et al. (1996), in a study of decision making of Australian armed forces, found that the structure of problems faced in the field are rarely routine. Eighty-seven percent of situations reported by participants contained some novelty. Thus, decision makers can expect to deal with ambiguous situations. This poses problems for both analytic and intuitive strategies. Analytic strategies require comprehensive and accurate data, which will generally not be available. Intuitive strategies rely on memory of past experiences, which may not be applicable to unfamiliar, ambiguous situations.

Intuitive strategies certainly are better suited to situations where time and data are limited (McMenamin, 1995). By using memory and avoiding an exhaustive evaluation of a large set of options, decision makers using an intuitive strategy can respond quickly. Also, memory-based strategies are less affected by the loss of external data than are analytic strategies. That is, decision making performance declines gradually as information becomes less available, in contrast to the catastrophic failures of analytic theories under this condition.

A potential drawback of intuitive theories is that they assume a large degree of expertise in decision makers. Not all decision makers are expert. In considering tactical decision making, it is important to support the less experienced naval officer, both in terms of general OR skills and specific mission-related knowledge. For example, an experienced but newly arrived ORO may have difficulty because he or she lacks knowledge about the mission, OR team, and ship.

Perhaps the key deciding factor between analytic and intuitive strategies is whether one can accept a merely workable solution or whether one requires an optimal solution. Decision makers in the real world, as we will see in detail later (Section 3.1.1), do rely on intuitive strategies more than analytic ones (e.g., Orasanu & Fisher, 1997, cited in Flin, 1998). The factors that affect this choice of strategy appear to be the time available, the level of risk, the situation complexity or familiarity, and the availability of information (Flin, 1998). It may be only under unusual circumstances, especially in naval C2, that decision makers have the time and data to engage in an analytic strategy.

Recalling Flin's (1998) classification of the three levels of command, analytic strategies seem a good fit for strategic command. Decision making at this level is long-term and there is strong incentive to find the optimal solution. Intuitive strategies seem better suited to the tactical and procedural levels. Here, there is greater time pressure and the goal is focused on accomplishing the mission in an effective manner more than finding the optimal way to accomplish the mission.

2.2.4 Heuristics and Biases

Another distinction relevant to decision making is between heuristics and algorithms. Analytic theories specify algorithmic solutions. That is, they contain explicit, step-by-step procedures for evaluating options and selecting the best COA. In contrast, heuristics are informal, intuitive strategies for solving problems (Sternberg, 1996, pp. 353-354). They specify simple steps, which are often based on probabilistic data, and are designed to work under a few general assumptions. Intuitive strategies are more heuristic-based, although the term heuristic is generally reserved for even less formal rules-of-thumb for making decisions (Colton & Gange, 1993).

Numerous heuristics have been identified and a thorough review is beyond the scope of this report. A few examples, however, are worth noting because they could be used in tactical decision making.

The *representativeness* heuristic is used to estimate the probability or likelihood of an event. One judges the probability of a event occurring based on how similar to, or representative of, the event is to a particular class of events (Colton & Gange, 1993; Kahneman & Tversky, 1972). For example, the more a radar track shares features with radar tracks of hostile aircraft, the greater the probability a commander might identify the track as a hostile aircraft.

The *availability* heuristic can also be used to estimate the probability of an event. Here, one judges the likelihood of an event based on how readily one can recall occurrences of the same event, or events of the same class, occurring in the past (Tversky & Kahneman, 1973). For example, if a commander sees track that indicates an aircraft on a direct course to the ship that is rapidly descending he may recall occurrences where that was prelude to an attack. In this case, the commander assigns a high probability that the aircraft is hostile.

Both of these heuristics are potentially useful, in particular because they are easy to apply and can be performed quickly. They both, however, could result in serious errors because they apply only under certain conditions. These heuristics ignore the *base rate*, or underlying frequency, of events in the world (Best, 1995; Kahneman & Tversky, 1972). Thus, in a given situation, hostile aircraft may be much less frequent than friendly and neutral craft. In this case, the representativeness and availability of hostile aircraft in the decision maker's mind would not themselves indicate the probability of a track corresponding to a hostile aircraft. Even if the track is consistent with a hostile aircraft, the relatively larger base rate of friendly and neutral aircraft would make it more likely that the aircraft is not hostile. The presence of a few highly salient features can potentially reduce a decision maker's attention to the base rate even further.

The main implication of heuristics for naval C2 is that decision makers can be biased and systematically distort the evidence considered. It is not that heuristics are inherently flawed

but that decision makers rarely test whether their underlying assumptions hold. In fact, decision makers are rarely aware of what those assumptions are.

Biases represent a big problem for decision making because they can render even an effective decision making strategy limited or inaccurate. Mayer (1989) discusses a number of constraints on decision making and problem solving:

- People systematically distort problems to be more consistent with prior knowledge by altering data, perceived constraints, and relations between problem elements.
- People focus on inappropriate aspects of problems; they may place needless constraints on possible solutions (often from previous problems) and are overly affected by past experience to determine what features to look for when representing problems.
- People change the problem representation during problem solving, often reformulating it inappropriately.
- People apply procedures rigidly and inappropriately, attempting to use past solutions when better solutions exist.
- People let their beliefs guide their approach to problem solving, often letting their beliefs affect the problem representation, strategies, solutions considered, and even which problems are attempted.

2.2.5 Analytic Models

This section reviews several analytic models that have been applied in the area of tactical C2.

2.2.5.1 Newell's Soar Framework

Perhaps the most comprehensive analytical model is *Soar* (Newell, 1990). It was developed as a unified theory of cognition and architecture for intelligent behaviour. That is, it provides principles and constraints based on cognitive processing (Kalus et al., 1996). Soar is based on the research of Newell and Simon, in particular their Logic Theorist (Newell & Simon, 1956, cited in Kalus et al., 1996) and General Problem Solver (Newell & Simon, 1961).

Soar is a general problem solver, intended to generate a COA for a problem in any domain by the process of problem space search. The general method of problem solving is by production systems that transform a problem space to reach a goal. In order to serve as a model of human cognition, the specific goals of Soar (Congdon & Laird, 1996, cited in Kalus et al., 1996) are to:

- Work on a full range of tasks from routine to difficult and open-ended.
- Represent and use all forms of knowledge, including procedural, declarative, and episodic.[1]
- Employ a full range of problem solving techniques.
- Interact with the outside world.
- Learn from the task.

[1] Three forms of memory are widely distinguished (see Tulving, 1983). Procedural memory consists of procedures and routines that are performed. Declarative memory, also called semantic memory, consists of factual knowledge about the world. Episodic memory consists of the store of personal memories related to events or episodes in one's own life.

In Soar, all problem solving is performed by selecting and applying operators to the current state (Kalus et al., 1996). In this way, the production system changes the current state to the goal state. Selection of operators consists of three actions:

- **Operator proposal**: an operator is proposed when the current conditions satisfy its pre-conditions for initiation indicating that the operator can be applied in that situation.
- **Operator comparison**: Soar uses preferences stored in memory to compare multiple operators and evaluate the set of candidate operators.
- **Operator selection**: an operator is chosen on the basis of preferences and is applied to the current state.

In the last step, operator selection, it is possible that no one operator is preferred. Soar can deal with this situation by randomly selecting an operator from the candidate set if all options are equally preferred. If there is no preference information, Soar must resolve the impasse. The process of resolving an impasse is not relevant for current purposes but it forms the basis of learning in Soar. Essentially, Soar is able to generate new productions and preferences by use of analytic processes to resolve an impasse and move the current state toward the goal.

Soar is a model of human cognition but also an architecture for creating intelligent agents (Kalus et al, 1996). How well it models human cognition, of course, depends on whether one accepts analytic views of decision making. Because Soar has been developed as an Artificial Intelligence (AI) approach to problem solving, it can be implemented as a computer system. Such a system can connected to sensors so that Soar operates on sensor input and produces outputs in the form of recommended COAs. This may explain the appeal of Soar and other analytic models, that they can serve as advisors to human decision makers, even if they do not function in the same way as human decision makers.

2.2.5.2 SHOR

Similar to Soar, the Stimulus – Hypothesis – Option – Response (SHOR) framework is another version of problem space search (Wohl, 1981, cited in Adelman, 1992). Although not as detailed as Soar, SHOR outlines for stages to decision making:

- Stimulus: gather/detect, filter/correlate, aggregate/display, and store/recall data.
- Hypothesis: create, evaluate, and select an hypothesis to explain the current situation.
- Option: generate, evaluate, and select a COA.
- Response: plan, organize, and execute the COA.

SHOR describes problem solving at a more general level than Soar but applies analytic strategies. The component processes of each stage are listed in Table 2.1 (see Webb et al, 1993). The stimulus stage identifies the problem space through the gathering and mental organization of information. Problems are solved by generating hypotheses about possible paths from the current to goal state. These multiple options are evaluated and the most promising is selected. The final step is to plan and execute a response to implement the selected option.

Much of the power of SHOR rests on the specific procedures for accomplishing each stage. This model can be implemented with a range of option analysis and evaluation

techniques (Adelman, 1992, pp. 12-14; Wohl, Serfaty, Entin, Deckert, & James, 1988) and is, perhaps, somewhat more pragmatic than Soar.

Elements/Processes	Functions
STIMULUS	Gather/direct Filter/correlate Aggregate Store/recall
HYPOTHESIS (perception alternatives)	Create Evaluate Select
OPTION (response alternatives)	Create Evaluate Select
RESPONSE Action	Plan Organize Execute

Table 2.1 - Evaluation Functions Inherent in Decision Making (Wohl 1981; cited in Webb et al., 1993)

2.2.5.3 Distributed Problem Solving

Another general analytic theory is the Distributed problem Solving Method (Van Daele & De Keysar, 1991, cited in Noakes et al., 1996). Again, the process of problem solving is one of moving through a problem space. It takes a subgoal analysis approach, however, in which the goal state is decomposed into a set of subgoals necessary to transform the initial state to the goal state. The subgoals are prioritized to order them in a temporal series.

A special feature of this model is that it applies to teams working on problems. Teams coordinate to decompose the task into subgoals then assign individuals to solve certain subgoals. Each individual then determines the strategies to be used in solving his or her subproblems. The model assumes little explicit coordination but a great deal of interaction to share information.

2.2.6 Intuitive Models

This section reviews several intuitive models that have been applied in the area of tactical C2.

2.2.6.1 Recognition Primed Decision-making (RPD) Model

One of the most influential models in recent years has been Klein's (1997) RPD model (see also, Klein et al., 1993; Leedom, Adelman & Murphy, 1998). Like all intuitive models, it eschews formal, logical processes and instead emphasizes recognitional and

pattern matching processes. The major principle of RPD is that the decision maker attempts to recognize the current situation and match it to a COA or solution that has previously been encountered.

Klein (1997) presents an updated version of RPD that incorporates additional processes meant to complement recognition. According to the complex RPD model, decision makers first appraise the situation in order to classify it as familiar or not, based on experience. The assessment of familiarity can be made in several ways, including:

- Recognition by matching features of the situation to prior events.
- Recognition of a whole pattern of features that fits a familiar story or scenario.
- Explicit recall of an analogy from another related domain.

If the decision maker is unable to recognize the current situation, the typical reaction is to seek more information. This sort of on-going situation assessment is a crucial component of decision making in RPD (Klein et al., 1993). Recent revisions to RPD (Klein, 1997) incorporate other means to resolve ambiguous situations. For example, decision makers can engage in active diagnostic processes such as *story building*. In this case, the decision maker deliberately notes features of the situation and attempts to create a detailed hypothesis or story that could explain that configuration of features. If there is more than one story compatible with the data, the decision maker can again gather more information and attempt to evaluate each story.

Once the decision maker has diagnosed the situation, he or she can use mental simulation to form expectations about future events. These expectations serve to test the working hypothesis, which, if disconfirmed, causes a decline in confidence in the hypothesis. This provides a means for the decision maker to monitor the situation and gauge the accuracy of memory before taking any action. If there are too many inconsistencies between the hypothesis and the situation, the decision maker must revise his or her hypothesis.

As the decision maker proceeds with mental simulation, he or she can also generate COAs. RPD assumes a satisficing criterion. That is, the decision maker does not attempt to generate multiple COAs but instead generates COAs in a serial fashion and sequentially evaluates each (Klein, 1992; Klein et al., 1995). Also, the decision maker does not use an analytic, feature-based method to quantify the value of the COA. Rather, the decision maker can make a holistic evaluation or rely on mental simulation of the consequences of the COA. Evaluation stops when the decision maker generates an acceptable COA.

2.2.6.2 Recognition/Metacognition Model

One potential flaw of intuitive theories is that they do not account for the active thought processes of human decision makers. One model that attempts to address this issue is the recognition/metacognition model (Cohen, Freeman, & Thompson, 1997). Like RPD, this model is based on people's good recognition skills but it also assigns a significant role to metacognitive processes. Metacognition refers to the active processes by which people monitor and control their other cognitive processes. Metacognition skills include:

- Identification of evidence-conclusion relations.
- Processes of critiquing that identify problems in arguments.

- Processes of correcting that respond to problems.
- "Quick test" processes that control when and how critiquing processes occur.

The recognition/metacognition model assumes many of the same recognition processes as RPD. That is, decision makers assess the situation, recognize a similar prior situation, and recall COAs that may serve as a solution to the current problem. This model, however, places greater emphasis on processes to test whether the recognized situation is a suitable hypothesis for the current situation.

According to the recognition/metacognition model, there is a four-step process, called STEP (for Story, Testing, Evaluation, and Plan), that is used to assess a working hypothesis and select a COA. In the *Story* phase, the decision maker attempts to construct an explanation or story that explains the current situation and predicts future events. In the *Test* phase, the decision maker compares these predictions to observations. This phase serves the additional purpose of guiding situation assessment so the decision maker can focus on the most relevant data. The *Evaluation* phase leads to a decision as to whether the Story is plausible. If not, the cycle begins again. Once a story is accepted, the decision maker enters the *Plan* phase and develops a COA. This process can be done alone or with a team.

2.2.6.3 Image Theory

Image Theory (Beach, 1990) is also consistent with RPD but places more emphasis on situation monitoring and incremental adjustment of a working plan. Unlike RPD, Image Theory does not focus just on the overall decision at hand but breaks the task into two sub-decisions. The first is a progress decision to determine whether any intervention or action need be taken. The second is a decision about which goals, plans, and COAs are needed to restore an acceptable situation. The latter decision is made only if the first reveals some threat or condition that requires action.

In making these two decisions, decision makers rely on recognitional and story building processes. Explanatory stories will indicate the consequences of acting or not acting in the situation. If action is needed, the story can be used to generate potential COAs and evaluate their effectiveness.

2.2.7 Hybrid Models

As the recent revisions to RPD indicate, there is growing recognition of the value of integrating cognitive and metacognitive processes with memorial and pattern recognition processes. Some models take this a step further and integrate analytic and intuitive concepts in order to take advantage of the strengths of each.

For example, Serfaty et al. (1997) has proposed a three-stage model. The stages define an intuitive framework for decision making. In the first stage, the decision maker retrieves relevant experiences and creates a schema of the situation. The second stage is to gather further data to fill in missing elements of the schema and create a more complete story. In the final stage, the decision maker uses the schema to recognize workable COAs. The analytic components are the processes by which the situation is explored and the COAs evaluated. The model assumes formal analytic processes to systematically explore and evaluate features of the situation. These processes are similar to logical hypothesis testing and confirms the truth of the hypothesis by the presence or absence of critical features. Rational comparison

methods are used to identify the best COA. The three-stage model contains levels of iteration not found in models like RPD, which emphasize rapid identification of a sufficient COA.

2.2.8 Applicability to Naval C2

A tremendous amount of effort has been devoted to developing models of human decision making. Unfortunately, no one model, nor any one class of models, best suits the tactical C2 domain. Instead, we can see valuable concepts in both analytic and intuitive models, as well as drawbacks.

The main advantage of analytic models is that they are highly explainable and prescriptive. The steps and processes involved in decision making are clearly specified. In fact, many of these models are implemented as computer programs. Consequently, these models are excellent tools for predicting and exploring human performance. In addition, the focus of analytic models is on obtaining the optimal solution to a problem. This, of course, is highly desirable in any domain as long as it can reasonably be achieved.

Analytic models apply very well in situations in which there are well-defined goals and in which factors and options can be explicitly defined. Tactical C2 is itself a formal domain with ROEs and generally well-defined responses permitted in any given situation.

A problem in applying an analytic model, however, is that the nature of the situation, and hence problem, may not be clear. In fact, it is possible that the goal state itself may not be unambiguous. This class of theory requires high quality data, which is rarely available in tactical situations. Without complete and unambiguous information, it is impossible to apply an analytic method of evaluation to COAs. In addition, analytic processes are computation intensive and as a result time consuming. It is unlikely that a commander will have the time to engage in a complete analysis of alternative COAs even if he or she has the capability of completing all the complex computations. Although obtaining an optimal solution is desirable, it is more important to act within the time constraints of the situation. Analytic models, as we will see when discussing empirical studies, do not seem to describe what people actually do when making decisions (see Section 3.1).

Intuitive models, in contrast, are very descriptive of human decision making (Klein, 1992, 1997). These models require much less computation and make use of pattern matching and recognition processes at which people are very skilled. As a result, these models are consistent with the need for rapid decision making in tactical C2. In addition, they are robust and can operate even when the available data is limited or ambiguous.

Although the emphasis is not on finding an optimal solution, the use of previous experience generally guarantees an acceptable solution. The decision maker can retrieve some COA that will resolve a threat or achieve a mission goal without any formal consideration of how that COA compares to other potential COAs. One drawback to this feature is that decision making is highly dependent on the familiarity of the situation. If the situation is novel or misrepresented, intuitive processes will not as helpful in generating and evaluating a new COA as would analytic processes.

One approach to resolving the limitations of analytic and intuitive models is to develop hybrid models that combine the best features of each. So far, fewer models of this sort have been developed. It is also unclear whether they do, in fact, combine just the best features of each

model. The three-stage model of Serfaty et al. (1997), for example, is highly computational while not ensuring an optimal solution.

A better approach may be to develop models of decision making specifically for the different levels of command; i.e., the strategic, tactical, and operational (Flin, 1998). Requirements can be specified more precisely for the individual levels than for an overall concept of command. Highly analytic models would be more appropriate to the strategic level than tactical or procedural. The tactical level, in contrast, would require more intuitive processes to allow speed under uncertainty.

The models considered in this section tend to treat decision making as an individual activity. This may reflect the theoretical and methodological background of the field in economics and psychology. From the naval perspective, decision making is a team process, even if one individual, the CO, bears ultimate authority and responsibility for determining the COA. The team nature of decision making has implications for the processes of decision making (i.e., how data is gathered and distributed, how alternatives generated and considered, etc.) and how factors such as experience and communication affect decision making performance. Team decision making will be discussed in Section 2.5.3 but it is an area that needs further theoretical development.

In addition to models of decision making, heuristics and biases must be considered. In general, people seem to use heuristics frequently. Rarely, however, do people check whether the necessary assumptions have been met. This leads to biases and errors in decision making. It seems likely that tactical decision makers are subject to the same heuristics and biases as decision makers in other domains.

2.2.9 Implications for Decision Support

Theories of decision making identify a number of specific tasks that decision makers perform (e.g., Rouse, 1980, pp. 119-135), including:

- Assessing the situation.
- Recognizing or retrieving similar experiences.
- Story building or hypothesis formation.
- Generating alternative COAs.
- Evaluating alternatives.
- Imagining consequences.
- Choosing a COA.
- Executing and monitoring a COA.
- Revising a COA.

All of these processes need to be supported to improve overall decision making. In particular, one must consider the hardware and software requirements for support in each area (Hollingshead & McGrath, 1995).

Decision support can be broadly defined as the provision of tools, information systems, computers, displays, and so on to help human decision makers. In practice, researchers typically refer to the concept of the Decision Support System (DSS) when discussing how to support decision makers (this concept will be discussed in detail in Section 3.2). A DSS is a

computerized system that processes information, computes recommendations, and presents information the its most useful form.

There are several approaches to designing decision support (Hair & Pickslay, 1992). One is to employ a computer system that implements analytical methods with formal validity but which are unrelated to the way humans make decisions. In this case, the DSS generates a recommended COA and is not intended to enhance a human decision maker's thinking. In this case, the DSS consists of an Expert System (ES) that engages in what designers consider to be the best decision making strategy (Hart, 1988). The value of this approach is that it takes advantage of the logic of analytic models and can generate an optimal solution. Expert systems can outperform human decision makers in both speed and accuracy and, thus, enhance decision outcomes (Glover, Prawitt, & Spilken, 1997). There are, however, drawbacks to this kind of DSS. First, the system may still require a large amount of computational time even if implemented by computer (Hair & Pickslay, 1992). Second, the system generates a COA and leaves it to the human operator to either accept or reject it. The operator has no access to the decision making process and must trust the system's accuracy, which can be difficult for experienced naval officers. The system may also induce passivity in the user and reduce vigilance (Glover et al., 1997). Finally, such systems are difficult for non-experts to maintain and use. An expert system approach is probably best suited to highly structured tasks in which there is a clearly delineated problem space (Manning, 1991).

A second approach is to not use the computer system as a reasoning tool but to manipulate how data are presented to a human user through the computer's interface. This approach is aimed at improving the human's decision making process rather than devising a formal solution for the user's consideration. The way information is presented, whether raw or synthesized, can have a large effect on the speed and accuracy of decision making (e.g., Jarvenpaa, 1989). The advantages of this approach is that it involves little extra computation time (it may, in fact, significantly reduce the user's mental effort) and retains all decision control with the human. The drawbacks are that the DSS can have negative effects on users' behaviour (Glover et al., 1997). For example, users may limit their decision making process to only that information given by the DSS. In addition, the system does not take advantage of the computer's computational power to further reduce the workload of the user. This approach is better suited to semi-structured domains where the nature of the problem space can vary from situation to situation (Manning, 1991).

A third approach is to design a DSS that makes decisions based on processes used by human decision makers (Hair & Pickslay, 1992). This can be termed the NDM approach. This approach shifts much of the computational burden from the human operator to the computer. However, because the system follows intuitive procedures it can be made transparent to the user so that he or she can follow the decision making process. This gives the user a basis for accepting or rejecting the system's output. Although the DSS is based on human decision making processes, it is built on a formal model and not subject to the heuristics and biases that affect human cognition.

The latter two approaches can be combined to create a computerized decision making model that highlights critical information and guides the user's decision making processes. Such a DSS serves less as a computerized advisor and more as a computerized assistant to tailor the informational workspace to the needs of the human decision maker and the demands of the task (see the concept of the embedded user model in Section 3.6.3).

It is also possible to combine an expert system with other DSS approaches. Sage (1986, cited in Hart, 1988) has proposed that hybrid systems are useful for supporting the user in assessing the situation, setting objectives for measuring the success of problem solving efforts, and generating alternative COAs. The system has a clearly advisory role but handles certain tasks that are computationally demanding.

From the three approaches, Hair and Pickslay (1992) recommend that several issues be addressed when designing a DSS. First, the DSS should be understandable by the users. A system that is "transparent" to the user (i.e., the user can identify and understand the computational steps of the system) helps guide the user's own decision making and is generally more acceptable and trustworthy to the user (Cochrane & Foley, 1991; Cyrus, 1991). Second, the DSS should not be seen as the decision making entity but rather as an aid to the human user. This is not to say that we should reject expert systems that produce recommendations but only that the user have access to the expert system's functioning. No system can be designed to deal with all possible situations, making it essential that users be able to modify the system to meet current needs. Third, the user needs access to the underlying knowledge base to understand the output of the DSS.

In addition, there are other issues important to naval C2 (Hollingshead & McGrath, 1995). First, decision makers act in teams, with each member responsible for specific tasks but dependent on other team members for information. The distribution of members over space, and perhaps time, adds complexity to the decision making process and the decision support needs. In particular, decision support design will have to consider communication and how and when information is passed to decision makers. Second, the design of previous C2 systems has affected how personnel make decisions in the past. Users have adapted to existing systems and developed strategies to use those systems effectively. Any development project will have to take into account how existing systems and practices limit decision making.

2.2.10 Recommendations

- Develop primarily intuitive models of decision making at the tactical and procedural levels of command and analytic models at the strategic level of command.
- Seek to integrate analytic and intuitive models of decision making to capitalize on the advantages of each model and too accommodate all levels of C2.
- Ensure models of decision making address certain specific factors, including time pressure, level of risk, situation familiarity, and quality of data.

2.3 Situation Awareness

SA is a major theoretical concept to intuitive theories. The term SA originated with pilots (Flach, 1996) and refers to the experience of fully understanding what is going on and being able to create a coherent "picture" of how the elements in the situation relate to one's goals. Thus, SA is a *state of knowledge* based on perception of the situation and understanding of the situation's meaning or significance (Endsley, 1997). Situation assessment, in contrast, is the *process* of acquiring SA. It consists of the perceptual and cognitive processes directed at acquiring information about the situation

and representing that information in a coherent mental model.[2] Both SA and situation assessment are seen as distinct from decision making, although they are highly related.

SA plays an important role in many complex tasks, such as flying, driving, and air traffic control (Endsley, 1995a). Air traffic controllers, for example, spend roughly 90% of their time processing information rather than implementing procedures (Kaempf & Orasanu, 1997). Tactical situations presumably place the same kind of information pressure on decision makers. The USS Stark incident, in which the crew of the Stark failed to respond to a threat, points out the necessity of an accurate mental model of events taking place for effective C2.

2.3.1 Endsley's Model of Situation Awareness

Endsley (1995a, 1997) has proposed, perhaps, the most comprehensive model of SA. The goal of her theory is to explain the nature of SA, how it is acquired, the factors that affect it, and the link between SA and actions (see also, Klein, 1989). Endsley's theory is diagrammed in Figure 2.1, which depicts the levels of SA, their relation to decision making, and factors affecting SA.

Endsley divides SA into three levels.

Level 1. This level deals with the perception of elements (situation variables) in the environment and consists of all the data that a person can gather about the situation..

Level 2. This level deals with the comprehension of the current situation. Thus, Level 2 SA consists of the synthesis of the disjointed information making up Level 1. As that information is related, the person interprets its significance with respect to goals. The result is a holistic picture of the situation, what is going on, and what it means.

Level 3. This level deals with the prediction of future conditions. Level 3 SA is a mental model of how the situation may project in the future and is achieved through knowledge of the dynamics of the elements of the situation.

Situation assessment, the process of acquiring SA, employs basic perceptual and memorial processes. Level 1 SA is built by perceptually sampling the environment. This step makes use of sensory and perceptual abilities and sensory and working memories. There is also a strategic component governing the search for information. Level 2 SA is a more cognitive process. The results of Level 1 are analyzed by comparing them to goal information in Long-Term Memory (LTM). This step involves creating a mental model of the current situation. Similarly, Level 3 SA involves manipulation of the mental model to predict future states.

SA is acquired over time. Information is continuously taken in and added to the ongoing understanding of the situation. Thus, SA is really a dynamic entity that can only be characterized for particular points in time. SA must itself also contain reference to time because knowledge of how elements are changing in the environment is often the most important component of SA (Endlsey, 1995a).

---

[2] These definitions of SA and situation assessment are not universally accepted. Some researchers make little distinction between SA and situation assessment (Cannon-Bowers & Bell, 1997; Garner & Assenmacher, 1997), whereas others view SA as the overall picture or judgment of the situation and situation assessment as the process of perceiving the situation.

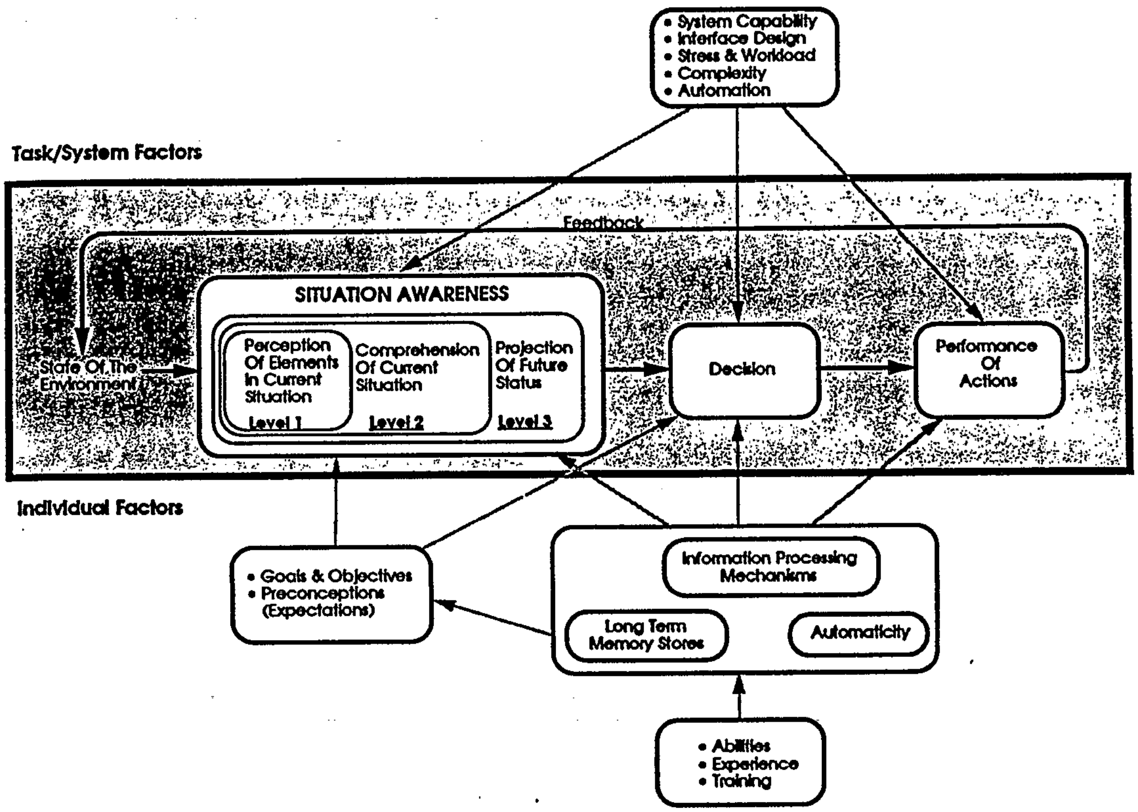

Figure 2.1 – Endsley's Three Stage Model of Situation Awareness (from Endsley, 1995a)

2.3.2 Factors Affecting SA

The power of Endsley's model is in its specification of the factors that govern SA. Endsley (1997) divides these factors into human properties and task and system factors. First, the following human properties affect SA.

Preattentive processing. Typically, this is a parallel process that detects and stores information in the sensory stores. Cue salience determines which portion of the environment is initially attended to, although higher order processes affect the perception of salience.

Attention. This constrains a person's ability to process multiple sources of information. It is the major limiting factor on Level 1 SA. Direct attention is needed to perceive and process cues to even a marginally meaningful level. Unfortunately, attention is also needed for concurrent decision and action processes causing potential problems of excessive mental workload. Information sampling is a strategy for managing the attentional limit (Wickens, 1992a; cited in Endsley, 1997). One attends to different sources of information in rapid sequence following a pattern established by priorities and frequencies of information change stored in LTM.

Perception. This is directed by the contents of Working Memory (WM) and LTM. Advance knowledge of the form, position, and characteristics of elements can facilitate the speed and accuracy of perception.

Working Memory. Perceived information is stored in WM and combined with existing knowledge to create a mental model of the situation that can be continuously updated as new information is taken in (Level 2 SA). WM is also used to project future conditions (Level 3 SA). Unfortunately, WM has a limited capacity and is the main limiting factor on Levels 2 and 3 SA.

Long-term Memory. This can be used to circumvent WM limitations to some extent. In particular, schemata and mental models stored in LTM can serve to guide attention and perception. Schemata provide frameworks for understanding information and organizing it in a coherent picture of the situation. Mental models are knowledge representations used to describe the operation of the system and predict its future operation.

Confidence level. People have some degree of confidence in their assessment of the situation. Their degree of confidence affects their decisions. Confidence serves a metacognitive role of helping a person direct his or her search of the environment, representation of the situation, and further goals.

Automaticity. Cognitive processes become increasingly automatic with practice. On one level, automaticity of perceptual processes seems beneficial to SA because automatic processes are faster and consume fewer mental resources than conscious, effortful processes. Automatic processes, however, tend to be closed-loop processes. Thus, automaticity poses a risk that a person will become less responsive to new stimuli and make limited use of feedback. Action slips, where a person initiates and performs an action without awareness, serve as an example (Norman, 1981). Action slips can be initiated in inappropriate situations or continued after appropriate conditions are terminated, resulting in errors and accidents.

Goals. These determine the perceptual search for information and selection of mental models and schemata. If the goals are inconsistent with the situation, a person may have trouble accurately appraising it.

Second, the following task and system factors affect SA:

System design. In complex tasks, a system intervenes between the human and the world. This system determines what information is available and how it is presented. Loss or distortion of information will obviously reduce the accuracy of SA. In addition, the way information is presented, or the way the user is required to interact with the system, can affect the attention, perception, and representation processes of the user.

Stress. Physical, social, and other forms of stress affect SA. A certain level of stress will actually enhance SA, leading to greater focus on elements critical to the task. Too much stress, however, consumes attentional resources and narrows the attentional field to a small number of elements. This creates the risk of building a misleading picture of the environment based on too little information.

Workload. The number of tasks or task components that a person must perform at any moment affects how much attention and effort can be devoted to perceiving and interpreting the situation. Complex tasks consume more resources than simpler tasks, leaving less for the task of acquiring SA.

Complexity. A complex situation requires greater analysis. Given limited cognitive resources, it will be harder to achieve a complete representation the greater the information content of the situation.

Automation. Automating the system can have positive and negative effects on SA. Automation gives the user more time to do situation assessment because the user is freed from manual tasks. Automation, however, can cause complacency and loss of vigilance if the user accepts that the system will do the bulk of the work. In addition, people sometimes fail to understand how the automated system works and use it improperly (Jones & Endsley, 1996). There is also a difference between being an active versus passive processor of information. Active exploration promotes greater comprehension and memory than passive viewing (e.g., Baddeley, 1990, Chs. 7-8).

2.3.3 Errors of Situation Awareness

One of the primary factors linked to failures in decision making is poor SA (Klein, etc). Endsley and colleagues (Endsley, 1997; Jones & Endsley, 1996) have developed a taxonomy of errors that can occur in SA and situation assessment. Their classification system highlights the problems that a person might have in developing a full understanding of the situation.

Errors at Level 1 centre around failures to perceive information. This can often result from some inherent lack of detectability or discriminability in the sensor data. Physical limitations prevent the operator from having complete access to information. Failure to perceive can also result from some failure of the perceptual and cognitive systems. In addition, an over abundance of data may overwhelm the person's perceptual capacity. Similarly, there may be a failure of data sampling due to a lack of an adequate strategy.

Errors at Level 2 centre around failures to comprehend the significance of events. Often novices do not possess a sophisticated enough mental model to interpret data and recognize its implications. Even an expert can select the wrong mental model from memory and mis-categorize data. If the person does not detect the discrepancy between the working mental model and the true situation, predictions based on the faulty SA will lead to errors.

Errors at Level 3 centre around difficulties in processing the dynamics of the situation. This failure may be due to limited cognitive resources or an inadequate mental model to predict future consequences.

2.3.4 Implications

Given that SA is an important component of intuitive decision making, it is important to seek ways in which complex systems can be designed to enhance SA. One way is by the design of the interface and controls (Flach, 1996). For any complex system, the interface should be designed to organize the data in a way consistent with the user's goals and needs for maintaining SA. This entails study of the nature of the task to identify what particular display elements are needed. In general, however, there needs to be greater emphasis on *data extraction* than *data availability*. Data extraction refers to the ability of the user to obtain information in the form needed, whereas data availability refers to the amount of data presented. Data extraction places an emphasis on the ease of use of information in a display. Displays can become disorganized fields of information if too much is presented or it is presented in a form that doesn't correspond to the user's needs (Flach, 1996). This places the

burden on the user to search for and transform the data to what is needed, increasing workload, decreasing perceptual accuracy, and requiring more time (Section 3.6.1 discusses related issues).

Endsley (1995a, 1997) offers some guidelines for interface design that emphasize these points. The unifying idea is that the designer must know and understand the user's needs for SA. In particular, Endsley (1995a) has identified eight interface features that should enhance SA based on her 3-level theory of SA:

1. Displays that present information in terms of Levels 2 and 3 SA will combat WM and attention limits.
2. Present information in terms of the user's major goals (Level 2 SA).
3. Make salient critical cues used to activate mental models and schemata.
4. Consider both top-down and bottom-up processing (in particular, reserve special design features for critical cues that indicate need for activating other goals).
5. Design to make global SA available at all times, with access to detailed info related to immediate goals as required; avoid displays that restrict access to SA elements at any time.
6. Filter extraneous information (not needed for SA) and reduce data by synthesizing and integrating SA elements into a form directly relevant to SA needs.
7. Provide system-generated support for projecting future events and states of system (Level 3 SA).
8. Support parallel processing of information to combat attention overload; e.g. by multi-modal presentation.

2.3.5 Recommendations

- Promote R&D to develop a better process model of SA that elaborates Endsley's model by specifying the cognitive operations involved in perceiving the situation, creating a mental model of the situation, and mentally projecting the situation into the future.
- Support Level 1 SA (perception) by reducing the amount of information presented to the user and making critical cues perceptually salient.
- Support Level 2 SA (comprehension) by helping the user to define situational goals, synthesizing information and relating it to the user's goals, and storing and presenting information needed to interpret data.
- Support Level 3 SA (projection) by highlighting salient changes in variables and providing tools for projecting current trends into the future.

2.4 Expertise and Training

The development of expertise - the high level of performance associated with practice and direct experience - is crucial for naval C2 teams. When defined behaviourally, expertise can be assessed in terms of the speed and accuracy of performance (Anderson, 1995, Ch. 9). A more useful definition, however, focuses on the cognitive processes that distinguish experts from novices. Researchers have sought to identify how memory and reasoning changes as a person develops expertise in a given domain. These changes indicate the kinds of mental processes that promote better performance.

2.4.1 Anderson's Theory of Expertise

A widely accepted theory of the development of expertise is Anderson's three-stage theory (Anderson, 1995). This theory describes the transitions in knowledge and skill level from novice to expert. It is a general model, meant to apply to cognitive and behavioural domains. Thus, it is, in principle, applicable to naval C2 but should be validated for any specific application.

According to Anderson's model (See Anderson, 1995, Ch. 5), the first stage of skill acquisition is the *cognitive stage*. At this stage, the individual develops declarative knowledge of the task, including its rules, procedures, constraints, and so on. All of this knowledge is factual and explicitly stored in the declarative memory store (Tulving, 1983). To access this knowledge, the individual consciously recalls facts about the task and attempts to use them to guide performance. Thus, at this stage, there is little procedural knowledge and the individual must devote a great deal of mental effort to directing his or her actions.

The second stage is the *associative stage*. In this stage, the individual gradually notices and corrects errors in his or her declarative knowledge and strengthens the associations between cues and problem elements and the actions or steps needed to perform the task. Thus, the individual forms greater associative memory (e.g., if-then rules) that can be used to guide behaviour. Performance is still not based on true procedural knowledge but the stored associations allow for faster, less effortful performance.

The final stage is the *autonomous stage*. Here, the associative productions developed earlier become automatic. Knowledge is stored in procedural memory that can be accessed implicitly and without conscious effort (or very little). Performance is initiated in response to situational cues that determine which learned procedures should be called upon.

This theory is consistent with observations of increased automaticity and proceduralism with expertise in a variety of domains (Anderson & Fincham, 1994). However, to fully understand expertise in cognitive domains (those requiring decision making and reasoning), we must identify differences in the knowledge structures of experts and novices.

Experts have more comprehensive and better organized domain knowledge than novices (Lesgold, Rubinson, Feltovich, Glaser, Klopfer & Wang, 1988; Sternberg, 1996; Wiley, 1998). This is almost a tautology but it indicates two important processes of expertise acquisition, namely the identification of crucial concepts, cues, and procedures in the domain and the organization of knowledge to facilitate rapid and accurate access to those concepts, cues, and procedures. Thus, as a person develops expertise, memory becomes specialized to recognize and retrieve appropriate responses to problems in that domain.

In addition to proceduralization, a process of schematization also occurs as one develops expertise. The expert creates mental frameworks that describe the kinds of problems encountered in the domain, the kinds of cues that indicate problems, and the kinds of solutions that can be applied. Anderson (1995) distinguishes between tactical and strategic learning. Tactical learning deals with the acquisition of skills and procedures, following the three stages of Anderson's model. Strategic learning refers to learning how to organize one's problem solving. Creating mental schemata to organize domain knowledge allows one to determine the best means to solve problems and quickly judge the applicability of solutions.

Anderson's theory implies that training should focus more on procedures than information. Because expertise involves acquiring sets of specific procedures, training can be enhanced by analyzing what those procedures are. Thus, a prerequisite to training is a thorough task analysis. The results of this analysis will be specific tasks and procedures that can be explicitly taught to novices. Such training will be especially effective if tasks are broken into components, each of which can be mastered more easily than the whole task in its entirety.

2.4.2 Theories of Training

Fallesen and Pound (1998) identify three broad approaches to training. One is a *formal approach* that is highly structured and focuses on explicit rules and procedures. The training is derived from formal theories and models and conveys a great deal of declarative knowledge. Lecture-based courses are examples of this approach.

A second approach is based on *intuitive* theories. This approach emphasizes performance, cue differentiation, and the acquisition of automaticity. Training is done by practical exercises with feedback to allow individuals to acquire experience comparable to being on the job.

The third approach is a *hybrid* approach that combines explicit teaching of concepts and procedures but reinforces this with practical exercises. The first step is to identify expert strategies for solving problems in the domain and teaching novices how to apply these strategies.

Fallesen and Pound surveyed the literature and identified over 65 individual training strategies (see Pound & Fallesen, 1994, cited in Fallesen & Pound, 1998). They found that informal training strategies are often preferred over formal strategies for training in military settings. That is, military training emphasizes recognitional and procedural techniques rather than informational content. The goal of training appears to be to help trainees develop experience that can be applied in the field, especially the ability to recognize situations. Thus, current training techniques are generally consistent with intuitive theories of decision making.

Nevertheless, Fallesen and Pound (1998) advocate a hybrid training approach. Instruction includes components of declarative description of C2 concepts and procedures. These are supported by exercises designed to reinforce those concepts and procedures. Trainees learn by applying skills in realistic scenarios. Thus, this approach involves more implicit learning than does classroom instruction but could be helpful early in training when novice trainees function in the cognitive stage (see Anderson, 1995).

Kozlowski et al. (1998) advocate a similar approach, involving practical exercises designed to give trainees experience in C2. To facilitate comprehension and retention, exercises start at a simple level and become more complex as trainees master the material. Kozlowski et al.'s (1998) approach also emphasizes the use of feedback to promote learning. In particular, they advocate the use of feedback to provide a) evaluation, b) attribution, and c) guidance.

Evaluation lets the trainee know whether his or her performance was correct or not. This is essential information but not sufficient to help the trainee learn. Feedback that provides attribution tells the trainee why his or her performance was incorrect. Specific *attribution* provides a breakdown of the correct and incorrect steps taken in an exercise and allows trainees to modify their understanding of the task at a detailed level. Finally, *guidance*

provides information on how to correct errors that trainees can incorporate in their understanding.

Radtke and Frey (1996) offer a list of 15 criteria for effective instruction, which are presented in Table 2.2. These criteria are consistent with other theories of training and the interactive, multimedia techniques discussed below. The criteria serve as specific goals for designing a training course in a complex domain such as naval C2.

Criteria for Effective Instruction
• Information is presented in an organized, logical order consistent with goals of instructions
• The presentation provides personalized feedback on the participant's performance
• The participant can review material as often as desired
• The participant can navigate in the presentation without becoming lost or stuck
• The participant sets the pace at which material is presented
• The presentation uses visual illustrations to make ideas and concepts clearer and more concrete
• The material is presented in concise and manageable "chunks"
• The learner can explore a topic of interest in more detail
• The presentation forces the learner to actively think about the material being presented
• The presentation holds the learner's attention by making the material interesting and entertaining
• The presentation provides opportunities for the learner to step back and think about how the small details fit together
• The presentation uses vivid visual images to help the learner remember important points or ideas
• The presentation uses animation and graphics to show ideas and processes that otherwise would never be seen in the real world
• Key points are presented in several different ways
• Important points are clearly identified with visual or sound cues

Table 2.2 – Criteria for Effective Instruction (from Radtke & Frey, 1996)

Both Fallesen and Pound's (1998) and Kozlowski et al.'s (1998) training approaches mesh well with current views on education. Many educators advocate a practical, learn-by-doing approach to teaching. The next section discusses some techniques that are useful in designing instruction in complex, procedural domains such as naval C2.

2.4.3 Training Techniques

This section reviews specific instructional techniques on more detail.

2.4.3.1 Learning-by-doing

Learning-by-doing, also referred to as experiential learning, encompasses a number of specific teaching techniques and strategies. The common element in all cases is an

emphasis on performance rather than acquisition of material. By encouraging trainees to do the tasks or jobs that they are being trained for, this technique helps trainees pick up the material as they apply concepts and procedures in exercises, case studies, and simulations. Learning-by-doing can often be implemented in a case study approach where trainees perform simulations of actual domain problems.

The principle of learning-by-doing is to establish a simplified model of the complex work domain (Vincent, 1998). Students learn procedures and concepts from the model then translate their knowledge to the real world. This technique establishes structural and process validity between the training and the real world task. In other words, trainees directly apply the procedures and material they are to master for the real world. This technique also enhances the "psychological reality" of the training by making trainees feel that they are already on the job in some respects.

There are numerous advantages to this instruction strategy. Learning-by-doing has been shown to (Dordick, 1997):

- Increase motivation and interest.
- Improve learning of the appropriate contexts in which to apply material.
- Produce better attentional focus on the task being learning rather than the task of learning itself (i.e. note taking, elaborating and integrating material).

Learning-by-doing techniques have been shown to improve ultimate performance levels after training in diverse domains, such as teaching science courses, economics and business administration (Galson & Oliker, 1976; Herz & Merz, 1998), report writing (Dordick, 1997), and medical diagnosis (Liu, Schneider, & Miyazaki, 1997).

Learning-by-doing is an effective way to combine teaching of conceptual and procedural knowledge, in a way that encourages people to integrate ideas and action.

2.4.3.2 Interactive Instruction

Interactive techniques can be likened to the mentoring (Casey, 1996) or tutoring (Merrill, Reiser, Merrill, & Landes, 1995) approaches. It enables learning-by-doing by providing constant direction and guidance.

A fundamental part of interactive instruction is that the trainee has some control of the direction and pace of instruction. Thus, the trainee indicates the sequence in which lesson elements are covered and how quickly they are covered. In this way, the trainee can use his or her self-assessment to determine those concepts least understood, the best way to integrate lesson elements with previous material, and the time needed to fully comprehend and incorporate new material. A course instructor must retain some control of the course in order to guide trainees. The instructor provides the organization and experience for trainees to chart a path through the course material.

A number of specific techniques can be used in interactive instruction. An important one is diagnosing the trainee's level of understanding. This entails providing accurate feedback that the trainee can use to guide how he or she interacts with the lessons. Also important are exercises to help the trainee identify his or her instructional objectives or the goals the trainee wishes to achieve (Lee, 1984). Students will not

necessarily have a clear idea of what they can or want to accomplish. By helping them to define goals, interactive instruction helps trainees build a mental framework to understand the material and its significance.

Interactive instruction places an emphasis on the use of questioning techniques (Lee, 1984). Questions, tests, and exercises, help refine learning objectives, provide the opportunity for feedback, and engage the trainee in active learning. Often, computer-based tutorials or simulations can provide excellent learning opportunities (Bostow, Kritch, & Tompkins, 1995). Questioning techniques can be combined with real-life examples and situations to provide even greater learning-by-doing.

Although interactive techniques emphasize trainee control, the role of the instructor as a support mechanism is crucial. Instructors should be used to initiate the learning process, answer questions, offer advice, and interpret the material and feedback given to trainees (Lee, 1984). Personal interaction enhances motivation and interest as well as the effectiveness of tutorials and tests (Bostow et al., 1995).

Interactive instruction has been shown to increase motivation and interest (Resnick, 1994) and performance (Lee, Rutecki, Whittier, Clarett, & Jarjoura, 1997) compared to traditional lecture instruction. In addition, interactive training improved performance in logistics command and control (Brecke & Garcia, 1995).

2.4.3.3 Feedback

A crucial component of interactive instruction is feedback. In fact, feedback and criticism are themselves important teaching strategies. For example, the act of recall has been shown to strengthen memory (Landauer & Bjork, 1978). Consequently, a potential learning strategy is to administer questions and exercises on a regular schedule. Repeated tests reinforce learned material and can help trainees integrate lesson elements to form a "big picture" of the course material.

To be effective, feedback should be specific, immediate, and constructive (e.g., Heward, 1997). By specific, we mean that feedback should not just give an overall evaluation – good or bad – but should indicate exactly what errors the trainee made, why they are errors, and how the errors can be corrected (Ziems & Neumann, 1997). In other words, feedback is valuable only if it explains the discrepancy between the trainee's understanding and the truth of the material.

Research has demonstrated that feedback is generally more effective if it is given right after a test (e.g., Kulik & Kulik, 1988). This helps the trainee make a direct link between the questions and the feedback and allows the trainee to effectively integrate the correct answer with the learned material.

Constructive feedback provides insight into where the trainee went wrong. It provides the correct answer but also suggests ways in which the trainee can improve his or her understanding. In this way, feedback not only tells the trainee how well he or she is doing but also offers guidance about how to further improve his or her understanding. For example, feedback should be used to assess current understanding and determine what new concepts are appropriate (Merrill et al., 1995).

2.4.3.4 Multimedia

Multimedia instruction is the combination of various media, such as text, video, audio, and graphics, for presenting material. It is simply the technique of making the best possible use of modern computer and audio-visual technology to expand the ways in which trainees can learn.

Multimedia is a natural choice when designing interactive, learn-by-doing instruction. These techniques rely heavily on real-world experience and detailed case studies and exercises. Multimedia offers a way to bring a great deal of real-world experience into the classroom. Combining video, audio, and other channels allows the instructor to simulate, with fair degree of fidelity, the work environment. Students are able to work on realistic problems while receiving realistic input.

Najjar (1998) reviewed the use of multimedia in education and found a great deal of research attesting to its effectiveness. Nevertheless, Najjar identified a number of principles to guide in development of multimedia instruction. There are several basic factors that must be considered in the design of multimedia courses: characteristics of the material, characteristic of the learner, and characteristics of the learning task.

Characteristics of the Material

Use the medium that best communicates the information. No one medium is the best and one must choose the particular media best suited to each element of the material. For example, auditory presentation is good for remembering small amounts of information for a short time but text promotes better long-term retention of large amounts of material. Pictures generally speed comprehension but limit the number of details that can be conveyed.

Use multimedia in a supportive not decorative way. Adding pictures and video merely to improve the appearance of a manual or course does nothing to improve learning. Each medium used must add something to the learning process that helps the trainee comprehend, review, elaborate, or enhance the material presented in another medium. For example, audio can be added to pictures to elaborate on details.

Use elaborative media. Some media, such as video and graphics, encourage elaborative processing. People generally find it easier to process visual information to a deeper and more meaningful level with less effort than text and auditory media. More importantly, combinations of visual and non-visual media can encourage even greater elaboration (Severin, 1967, cited in Najjar, 1998).

Make the interface interactive. As discussed previously, interactive instruction promotes better learning and memory by encouraging elaborative and integrative processing of material. The trainee has control and can actively explore the material. Interactive interfaces can have a positive effect on learning from multimedia because active exploration facilitates the integration of new material with existing knowledge (see Najjar, 1998).

Characteristics of the learner

Use multimedia with naïve learners. Students with little prior exposure to the material are likely to benefit most from multimedia. It provides naïve learners with a better framework for bringing together all the material.

Use multimedia with motivated learners. Multimedia can be more cognitively challenging because the trainee must take a more active role. Thus, motivated trainees tend to do better. There are, however, ways to increase motivation through the design of multimedia instruction. Relating the material to the learning objectives of the trainees enhances interest and motivation. Likewise, providing immediate, corrective feedback improves motivation by making trainees aware of how well they are progressing.

Use multimedia with adult learners. Adults have the requisite cognitive abilities to deal with multimedia and the active learning strategies that it promotes. Adults are generally better able to integrate information from several simultaneous sources.

Characteristics of the learning task

Use multimedia to focus the learner's attention. Learning involves prolonged concentration on the course material, which can be difficult to maintain even for adults. Graphics, video, and other techniques such as humor and story-telling can help trainees maintain their attention. This does not guarantee success but it is a necessary first step.

Encourage learners to actively process the material. Combined with interactive instruction techniques, multimedia can present trainees with "learning challenges," or problems that require them to explore and manipulate the material. This sort of active processing promotes better learning than passive receiving of information. Simultaneous presentation of information over two or more media can promote better integration of material than sequential presentations.

2.4.4 Applying Training in Team Situations

Part of the naval C2 domain is working in teams. Thus, training should contain some components devoted to teamwork and necessary team interactions. Following Cannon-Bowers et al. (1995) theory of team competencies, C2 teams should receive some training in realistic settings and in intact teams to promote learning of task-contingent and team-contingent competencies.

More specifically, learning in the field or realistic simulators promotes both familiarization with tasks and equipment but also skills in performing tasks with the equipment (Johnson & Cannon-Bowers, 1996). In addition, Johnson and Cannon-Bowers (1996) advocate a set of team-oriented training concepts.

Stress exposure training. Naval C2 is a high stress job and trainees need training in techniques to mitigate the effects of stress. These techniques include information about the typical physiological and mental reactions to stress and ways to use decision support systems to reduce workload and maintain team functioning.

Cross-training. By learning about the tasks and responsibilities of other team members, trainees develop shared mental models of the C2 task. This improves aspects of team performance, such as communication and subtask assignment.

Team adaptation and coordination training. Naval C2 demands that teams adapt to frequent changes in the tactical situation. Teams should practice together ways to respond to various tactical situations and coordinate communications and responses.

2.4.5 Recommendations

- Promote R&D to identify training techniques (especially experiential) that promote the development of expertise specifically for each OR position of the CPF.
- Conduct further development of performance-based criteria for effective training specifically for each OR position of the CPF
- Include individual and team training components in any training program.
- Validate Anderson's stage model of expertise for the CPF OR and use it to identify appropriate training techniques for each stage of expertise.

2.5 Teams

Naval C2 decisions are made by teams. A single commander may issue an order and have responsibility for the actions of the ship but the decision making process, from data gathering, to interpretation, to COA generation, is performed by a team. Unlike a group, which is a collection of individuals working according to their individual agendas, a team is structured entity. A team can be defined as "two or more interdependent individuals performing coordinated tasks toward specific task goals" (Nieva et al, 1978, cited in Fleishman & Zaccaro, 1992). Thus, C2 operations depend not only on individuals but individuals linked and organized in specific ways (i.e., in a CT).

Probably the most important aspect of teams is that team members are interdependent. Teams are used when a task cannot be accomplished by an individual or a set of individuals working on their own (Cannon-Bowers, Oser, & Flanagan, 1992). Whatever a particular member is doing, he or she in some way needs the other members. This affects the way individuals perform their duties and it affects the way the team as a whole performs.

Another important aspect of many teams is that they are hierarchically structured. Members differ in status or responsibility for team performance (Hollenbeck et al., 1995). This is true in naval teams, which will have a clearly defined chain of command, with one person who bears ultimate responsibility. In addition, team members have different roles or functions that they perform within the context of team functioning. Members bring different expertise, knowledge, and information to their functions. Thus, teams have what can be called distributed expertise (Hollenbeck et al., 1995). That is, the team as a whole can exhibit behaviours and expertise that no single member could.

The interdependent nature of teams creates teamwork needs and behaviours not relevant when considering individual performance. These are behaviours that help maintain team coordination and promote effective communication between members. The following is a list of the most significant team behaviours (Cannon-Bowers et al., 1992; Bowers, Braun, Morgan, 1997):

- Coordination.
- Mutual adjustment.

- Compensatory behaviour/adaptability.
- Communication.
- Cohesion.
- Leadership.
- Team situation awareness.

A portion of what a team does, then, is aimed at simply maintaining the integrity and proper functioning of the team itself. When theorizing about decision making by teams, these are areas that must be considered in addition to the perceptual and cognitive abilities of individuals.

2.5.1 Theories of Teamwork

This section reviews several theoretical perspectives on teamwork.

2.5.1.1 Task and Organization Structures

One way to analyze team performance is to divide what a team does into two categories; task-oriented and organization oriented (Sengupta et al., 1996).

The task structure is the set of tasks that the team must perform and the rules, constraints, and relations that govern how tasks may be performed. Sengupta et al. (1996) argue that team performance can be predicted, in part, on the basis of characteristics of the tasks because the task structure governs what the team strives to do.

In particular, a task can be defined on several dimensions:

- **Uncertainty**: the extent to which the state of the situation and task is not known (e.g., because data is limited or unreliable).
- **Time**: the tempo of the task and the urgency of action.
- **Complexity**: the extent to which the task contains six characteristics - multiple attributes, multiple solution paths, multiple possible outcomes, conflicts between outcomes, correlations among cues, and uncertain or probabilistic links between cues.
- **Resources**: the physical materials needed to do the task.
- **Information**: the data needed to do the task.
- **Task coordination requirements**: the interdependencies among task activities, such as balancing resources and information, and sequencing and controlling actions.

These components of performing a task apply to both teams and individuals. In addition, teams must be analyzed in terms of their organization structures, which are the factors governing how the team maintains itself as a team. These include:

- **Team topology**: the hierarchical organization and authority/responsibility relationships between members.
- **Facilities and infrastructure**: the means by which information and resources are distributed through the team.
- **Communications links**: the connectivity (who communicates with who) and media richness of communication.
- **Distribution of activities**: the manner in which sub-tasks are assigned with respect to the distribution of expertise and knowledge across the team.

Sengupta et al.'s (1996) classification scheme describes the dimensions along which teams can be measured and indicates potential theoretic dimensions of teams. The scheme itself, however, does not explain how teams manage their task and organization structures.

2.5.1.2 Team Competencies

Cannon-Bowers, Tannebaum, Salas, & Volpe (1995) offer a classification scheme that address the kinds of skills that teams need in order to perform task and organization functions effectively. They developed a classification of team competencies, which consist of the knowledge underlying a team's performance, the skills needed to perform a task, and the appropriate attitudes on the part of team members.

Cannon-Bowers et al. (1995) identified lists of specific competencies in all three categories based on a survey of the team research literature. These are listed in Table 2.3. Although no single team will possess all of these competencies, they are Knowledge, Skills, and Attitudes (KSAs) that can contribute to effective team performance. The thrust of Cannon-Bowers et al.'s (1995) classification scheme is to determine the kinds of competencies different teams can and should possess.

Knowledge Competencies	Skill Competencies	Attitude Competencies
• Accurate, shared mental models	• Adaptability	• Positive attitude toward teamwork
• Understanding of the nature of teamwork	• Shared situation awareness	• Positive attitude toward team concept
• Knowledge of overall team goals, missions, and objectives	• Performance monitoring and feedback	• Positive attitude toward collective efficacy
• Knowledge of boundary spanning	• Interpersonal skills	• Cohesion (attitude toward developing and maintaining interpersonal relations within the team)
• Knowledge of other team members' roles	• Communication skills	• Mutual trust (attitude toward team)
• Cue-strategy associations	• Decision-making skills	• Shared vision (attitude regarding direction, goals, and mission of the team)
• Interpositional knowledge	• Coordination (ability to work together, anticipate each other's needs, inspire confidence, etc.)	
	• Cooperation (ability to compensate for others' weaknesses, offer help only when needed, pace activities, etc.)	

Table 2.3 – Team Competencies
(from Cannon-Bowers et al., 1995)

To this end, Cannon-Bowers et al. (1995) divided team competencies according to two criteria. The first is whether the team competencies are specific to a particular team. In this case, the KSAs exhibited by a team were developed within the context of the particular team members and cannot be taken by individuals to new teams. Team-generic competencies can be held by individual members and influence team performance regardless of the teammates involved. Examples include communication skills, leadership, and attitudes toward teamwork. Team-specific competencies relate only to specific teammates; e.g., knowledge of teammates' characteristics, compensation strategies for teammates, and team cohesion. The second criterion is whether the team competencies are specific to a particular task. In this case, the competencies relate to a given task and how it is performed and cannot be used, even by the same team, for other tasks.

Based on these criteria, Cannon-Bowers et al. (1995) developed a two-by-two contingency table, shown in Figure 2.2 below. The contingency table divides the kinds of competencies that teams can possess into four categories:

- Context-driven competencies, which depend on a specific team and a specific task.
- Team-contingent, which are specific to a particular team but generic with respect to tasks.
- Task-contingent, which are specific to a given task but can apply to a variety of teams.
- Transportable, which are not specific to either team or task.

Teams can be analyzed with respect to this classification scheme to determine whether they possess the requisite competencies for effective team performance. Which specific competencies are needed depends on the natures of the team and task. Naval tactical C2 can be classified as context-driven and/or task-contingent. Certainly, performance depends on the unique nature of the C2 task. It may also depend on the specific team because team members must work very closely with one another and develop rapport. A high degree of interdependency among team members and a stable task environment indicate a context-driven team (Cannon-Bowers et al., 1995). C2 teams, however, may not be entirely context-driven because there is a fair degree of turnover in personnel aboard a ship and teams must maintain effectiveness when integrating new members. When turnover is rapid, task-specific competencies are required but team-specific competencies are less critical (Cannon-Bowers et al., 1995).

		Relation to Task	
		Specific	*Generic*
Relation to Team	*Specific*	Context-driven	Team-contingent
	Generic	Task-contingent	Transportable

Figure 2.2 – Types of Team Competencies (from Cannon-Bowers et al., 1995)

The requisite competencies for context-driven teams are listed in Table 2.4. They include competencies that will be specific to the task as well as competencies developed

with respect to specific team members. The more the team is able to adapt its KSAs to the task and team, the more effective it will be.

The requisite competencies for task-contingent teams are also listed in Table 2.4. There is a fair degree of overlap with those of context-driven teams because the focus is again on adapting to the specific task. In contrast, however, task-contingent teams require certain generic team competencies that allow individuals to transport their KSAs and interact effectively with others regardless of the team they are in.

Context Driven Teams	**Task-Contingent Teams**
Knowledge Competencies	**Knowledge Competencies**
• Accurate knowledge of one another • Knowledge of special responsibilities in the team • Common task models • Task-specific information flow • Accurate knowledge of cue-strategy associations • Common understanding of team goal and missions	• Accurate models of task and problem • Understanding of task-specific roles • Requirements for task sequencing • Team-role interaction patterns • Mechanisms and procedures for task accomplishment
Skill Competencies	**Skill Competencies**
• Task organization • Balancing workloads based on task and team demands • Shared problem solving models • Flexibility/adaptiveness	• Leadership or team management • Feedback and performance monitoring • Coordination • Assertiveness • Planning • Situation assessment
Attitude Competencies	**Attitude Competencies**
• Collective efficacy	• Task-specific attitudes towards teamwork • Attitudes toward working as a team

Table 2.4 – Context Driven and Task-Contingent Team Competencies (from Cannon-Bowers et al., 1995)

The team competencies identified serve as a taxonomy for decision support in team environments. They also lead to a number of proposals for team training requirements (Cannon-Bowers et al., 1995). These proposals are listed in Table 2.5 but are summarized briefly here. For the purposes of training, C2 teams should be treated as both task-contingent and context-driven teams. That is, training must deal with those

competencies specific to the task because the C2 environment poses stable problems for teams to solve. Training cannot focus as exclusively on team-specific competencies because there will be turnover in team composition. Nevertheless, the highly interdependent nature of C2 requires some focus on the particular team itself.

Proposition	
1	High interdependency in a team task requires team members to possess context-driven competencies.
2	Teams that operate in an environment that is fairly stable require task-specific but not necessarily team-specific competencies.
3	In teams where turnover is rapid, task-specific competencies are required and team-specific competencies are less crucial.
4	Team members who hold membership in multiple teams require, at the minimum, transportable team competencies.
5	When team members interact together across a variety of tasks, team-specific competencies are required; task-specific competencies may be less feasible (or necessary) to develop in such cases.
6	Teams that require team-specific competencies, whether they fall into the team-contingent or context-driven categories, will benefit from training as intact teams.
7	Teams that require task-specific competencies, whether they fall into the task-contingent or context-driven categories, should be allowed to practice in the actual task environment (or in one as close as possible).
8	Training for teams that require team-specific competencies, in either the context-driven or team-contingent categories, should incorporate feedback that leads to shared or common expectations for task performance.
9	When transportable competencies are required, some training can be focused at the individual level.
10	Task simulation may be an effective training strategy for teams that require task-specific competencies requiring actual practice. Further, task simulation can be an effective means of imparting team-contingent competencies if the operational team members are allowed to practice together (and only under these conditions).
11	Cross-training may be effective for teams that require exposure to the task (that is, task-specific competencies, whether they fall into the context-driven or task-contingent categories).
12	Positional knowledge training may be useful for teams with task-specific competency requirements, either context-driven or task-contingent.
13	Training to impart context-driven competencies should include guided practice that exposes the actual team members to the variety of situations they may confront on the job. When the actual team cannot be trained intact, guided practice may be useful as a means of training task-specific (but not team-specific) competencies.
14	Lecture-based training may be appropriate for transportable competencies but should be considered only as a first step for other types of competencies, since these require experience with the actual task or team.
15	Role playing may be used effectively to training team-contingent competencies when it involves the actual (operational) team.
16	Passive demonstrations of the task may be an effective means of training task-contingent competencies.

Table 2.5 – Propositions for Team Training (from Cannon-Bowers et al., 1995)

To facilitate learning of task-specific competencies, teams should practice is the actual task environment, or at least a high fidelity simulation. Effort should be put into achieving realism to promote direct experience. Task simulation can be an effective training technique if the simulation exposes team members to the variety of situations they will confront on the job. In addition, guided practice with plenty of opportunities for feedback helps team members learn their specific roles within the team.

To facilitate learning of team-specific competencies, teams should train as intact teams. That is, teams should be formed prior to training and progress together through the training program. This allows team members to learn and understand the characteristics, strengths, and weaknesses of teammates. It also promotes a shared representation of the task and team. For C2 teams, however, this should be balanced against the need to develop team-generic skills. Thus, some training should focus on the individual and development of competencies related to the individual's role in the generic team.

2.5.1.3 Team Functions

Other theories have examined in greater detail the processes by which teams perform the teamwork component of team performance. Nieva et al. (1978, cited in Fleishman & Zaccaro, 1992), for example, proposed four general kinds of team functions: team orientation, team organization, team adaptation, and team motivation. These functions help the team maintain cohesion and perform the team task more effectively.

Shiflett et al. (1982; cited in Fleishman & Vaccaro, 1992) refined this analysis and developed five classes of team functions. The first is *team orientation*, which describes activities related to exchanging information between team members and assigning sub-tasks. The second is *resource distribution*, which entails assignment of physical and informational resources to particular team members to match resources to task functions. The third consists of *timing functions*, governing the sequencing and regulation of team tasks. The fourth is *response coordination*, which includes actions to coordinate the outputs of individual team members and generate a team response. The fifth consists of *motivational activities*, involving the development of team norms and rewards and resolving internal conflicts (see also Nieva et al., 1978, cited in Fleishman & Vaccaro, 1992).

Fleishman and Vaccaro (1992) recommend adding two other functions to this model. The first is *system monitoring*, which involves activities aimed at detecting and responding to errors in other team processes. Team members must devote some effort to "watching out" for their teammates to ensure good teamwork. The second is *procedure maintenance*, which is the monitoring of individual and team behaviour to ensure that established performance standards are met. Thus, individual team members must take on the responsibility of evaluating the performance of the team as a whole.

The value of these team function models is that they indicate processes and actions necessary for teams to work properly. In particular, they indicate a need for some explicit training in team management. Coordinating team communication, resource distribution, timing, and other functions become increasingly difficult under high workload and stress. An experienced team may have developed good operating procedures in these areas but their procedures may not be automatic. In addition, team functions provide a basis for developing measures of team performance.

2.5.2 Factors Affecting Team Performance

2.5.2.1 Coordination

The ability of team members to pass information, divide tasks, sequence task activities, and combine the results of subtasks affects both taskwork and team work. Coordinating task functions is a bottom-up process, driven by the demands of the task (Bowers et al., 1997). In contrast, coordination of team functions is a top-down process related to maintaining team cohesion and effectiveness.

2.5.2.2 Communication

Research indicates that team performance is inversely proportional to the amount of communication within the team (Bowers et al., 1997). One reason for this may be that as teams spend more time communicating, they can spend less time on the task. In addition, not all communication is of equal value. Vague or inaccurate communication impairs the ability the team to complete its tasks. Because communication itself imposes workload, each instance of communication must convey maximal amounts of pertinent information.

2.5.2.3 Workload

Individual performance is not a strictly linear function of workload. That is, performance degrades at both low and high levels of workload because available cognitive resources for a task can be restricted by both boredom and excessive competing demands. Team workload is partially a function of individual workload (Bowers, Braun, & Morgan, 1997). Consequently, when individual members have too much or too little to do, the team's performance will suffer because their subtasks will not be performed effectively. In addition, however, team workload is a function of the effort needed to coordinate and manage the team. Inefficiency in the division of labor will add to the workload of individual team members.

2.5.3 Theories of Team Decision Making

Although a good deal of research has focused on how teams perform teamwork functions, less has considered how teams perform other functions, such as decision making. Team decision making consists of any decision making activity involving more than one person. This need not be collaborative, as in groups reaching a consensus on a COA, but also includes hierarchically organized teams where a single leader issues directives but relies on other team members to gather data and offer recommendations.

Theories of team decision making have often been based on extensions of theories of individual decision making (Ilgen et al., 1993). Generally, theories of individual decision making share two basic premises. The first is that individuals base their decisions on some finite set of cues or data points. The second is that individuals combine the cues according to some procedure to reach a decision. As we saw earlier, there are numerous analytic and intuitive strategies to do this. Extending these kinds of theories to teams raises issues. In teams, access to cues is distributed across members, and often no single member has access to all cues. Similarly, the procedural steps in decision making are distributed across the team so that a number of individuals pass information and perform a portion of the overall decision making process. Thus, team decision making theories must incorporate procedures by which

team members can interact in data transmission and summing of the results of individual computations.

Hollenbeck et al. (1995) presents a model that extends an individual decision making theory to the team domain. They base their model on Brunswick's (1940, 1956, cited in Hollenbeck et al., 1995) *lens model*. This is an analytic model that focuses on evaluation of cue-outcome associations. In the lens model, the decision maker assigns weights to each cue on the basis of its usefulness in generating a correct decision. The decision maker then determines the weighted sum of all available cues to determine the decision outcome.

To extend the lens model to the team domain, Hollenbeck et al. (1995) added a hierarchical structure to cue evaluation. Individual team members interact with unique but partially overlapping subsets of the available cues. Each member then generates a recommendation by evaluating the weighted sum of cues. A team member at the next higher level, perhaps the team leader in a small team, evaluates the recommendations. This evaluation is done in a fashion similar to cue evaluation. The leader assigns weights to each team member reflecting the level of trust, task relevance, and communication validity that the leader has in each member. The leader then computes a weighted sum to determine which recommendation to accept and passes it on to a higher level, if applicable. Thus, recommendations are propagated through the hierarchy until a single leader makes the final decision.

The area of team decision making seems theoretically under-developed. Hollenbeck et al.'s (1995) model is one example but there are many more theories of individual decision making. In particular, intuitive theories have gained a great deal of importance. Work needs to be done in bringing intuitive concepts into team decision making theories.

2.5.4 Team Situation Awareness

One theoretical construct that must be integrated with theories of team decision making is team SA. This construct is crucial for individual decision makers and must play some role in the performance of teams. Unfortunately, defining team SA is not necessarily a straightforward matter. Team SA refers to shared knowledge of the situation. Given a set of individuals with their own SA, team SA arises from the overlap of individuals' SA. How this overlap should be used in determining team SA, however, is a subject of debate.

One view is to define team SA as the area of informational overlap common to all team members (Noakes et al., 1996; Elliott, Schiflett, & Dalrymple, 1996). Thus, team SA would consist of the knowledge possessed by each individual. In this definition, members can possess their own individual SA and at the same time have an understanding of the shared team SA (Cannon-Bowers et al., 1993). Individual SA may guide performance of subtasks but team SA guides interactions between members.

Another view is to assess team SA based on the extent to which each member has the knowledge he or she needs to perform his or her duties (e.g., Endsley, 1995a). In this case, members need not have overlapping SA in order for the team as a whole to possess team SA. If every team member has the appropriate SA for his or her requirements, the team as a whole can perform well. If individual members lack the SA necessary for their roles, the team will suffer. In this view, it is theoretically possible for a team to have good team SA even if the members possess completely disjoint sets of knowledge (assuming the team and tasks

structures require no overlap in responsibilities). This view treats team SA as a more abstract concept consisting of knowledge distributed across the individual members.

A third view considers team SA to consist of a shared mental model of the situation (Rouse et al., 1992). This view is similar to the first but treats team SA as an active construct shared and maintained by the team rather than the knowledge that happens to be common to all members. In this view, individual members share information, which is a component of team SA. In addition, however, team members want to understanding the situation and predict what will happen in the future. To this effort, members coordinate their own mental models to achieve a common representation of the situation. Although each member must hold a mental model in mind, this view assumes explicit communication processes that members use to revise their own SA to be more like that of all other members.

2.5.5 Characteristics of Effective Teams

In general, effective teams will be ones that can balance taskwork and teamwork functions under the constraints of limited time, mental resources, and communications bandwidth. Hackman (1987, cited in Bowers, Braun, & Morgan, 1997) offers six more specific teamwork behaviours that characterize effective teams:

- Consider the competencies of individual team members.
- Encourage individual members to learn from one another.
- Foster creative planning.
- Maximize the efficiency of performance strategies.
- Maintain coordination and motivation.
- Create a sense of team spirit.

In addition, Cannon-Bowers et al.'s (1995) classification of team competencies indicate specific abilities or traits necessary for effective teams. The lists of competencies in Table 2.3 serve as descriptors of effective C2 teams. In particular, effective C2 teams require knowledge of the task and individual roles and skills in sharing information and dividing workload.

2.5.6 Recommendations

- Promote R&D to develop a comprehensive theory of team decision making that accounts for team SA, communication, and generation of recommendations and COAs.
- Apply Cannon-Bowers et al.'s (1995) team competency framework to identify specific competencies required by naval C2 teams.
- Conduct further study of the characteristics of effective C2 teams to develop a set of desired KSAs.

2.6 Human-Computer Interaction

Naval C2 depends on complex systems and computers. Thus, an important determinant of the effectiveness of decision makers is the interface between the operator and machine. Human-Computer Interaction (HCI) is a rapidly growing field of study but it remains largely practical rather than theoretical. Unfortunately, empirical answers in this field are limited (see Howell et al., 1993). Most

success has been in developing guidelines for narrowly defined tasks common to all computer users, such as text editing. Consequently, there are few absolute standards.

Nevertheless, researchers have identified major design issues and developed some general principles to guide specific HCI design problems. These issues centre on generic supervisory control (Rouse, 1985; Sheridan, 1987, cited in Howell et al., 1993), which is a high-level concept of all HCI.

2.6.1 Cognitive Fit

Perhaps the most general principle of HCI is that of *cognitive fit* (Vessey, 1991). Each domain or task establishes constraints on how the operator can use the system. These constraints arise from the information needed to perform the task, the kinds of strategies that will be effective, and perhaps computational limitations of the system. The goal of the operator is to use the most effective strategies as implemented on the system. Thus, the goal of the interface should be to make it possible and easy to do this. Cognitive fit describes the correspondence between the representation of the task in the system interface and the operator's mental representation. In other words, there must be correspondence between the way the system organizes and presents information and the operator's information needs, between the kinds of manipulations the system supports and the kinds of manipulations the operator wishes to perform, and between the way the system collects input from the operator and the kinds of responses the operator needs to make. When there is a mismatch, the operator is blocked from using the best strategies for doing the job. It may still be possible to accomplish the task but it is more difficult, increasing the operator's mental workload and frustration, and increasing the time needed to perform the task and the errors made (e.g., Jarvenpaa, 1989; Johnson, Payne, & Bettman, 1988, cited in Adelman et al., 1993).

2.6.2 Ecological Interface Design

Design of HCI requires attention to three levels of cognitive processes, skill-based, rule-based, and knowledge-based (Howell et al., 1988). The skill-based level deals with the physical interactions of the operator with the system. The rule-based level deals with the compliance of the operator with rules and operating procedures. The knowledge-based level deals with diagnostic and inference processes and the strategies used to solve problems. All three levels are important and interdependent.

Vicente and Rasmussen (1992) have developed the Ecological Interface Design paradigm to address HCI issues at all these levels. The premise of this approach is to analyze systems at different levels of abstraction, ranging from structural to functional. Descriptions at the various levels of abstraction provide answers to different issues that arise when considering operators' skill-based, rule-based, and knowledge-based interactions with the system.

According to Ecological Interface Design, skill-based behaviour is supported by an interface that is consistent with the movements required by the task. The mapping between perception of information and action is critical. Thus, interfaces should be designed so that elementary movements can be grouped or chunked into more complex routines that can be cued by visual features of the interface.

Rule-based behaviour is supported by a consistent one-to-one mapping between the constraints of the task and the cues or signs provided by the interface. Thus, there must be consistency in the way the interface displays information so that it can elicit consistent and appropriate

actions from the operator. For example, the interface should present the operator with cues that can be used to select an appropriate action from a menu. The goal is to avoid "procedural traps" in which the operator initiates an inappropriate action because the interface cues were ambiguous.

Knowledge-based behaviour is supported by provision of design features that facilitate goal-direct behaviours. At this level, the interface should contain functions that support the operator's strategies for performing the task. Information in the interface display should reveal the problem space to the user in a clear and unambiguous format. In other words, the display should serve as an externalized mental model of the task that the operator can manipulate.

2.6.3 Recommendations

- Analyze the task requirements at the skill-based, rule-based, and knowledge-based levels.
- Design for cognitive fit by undertaking a thorough analysis of the task, identifying goals, information requirements, procedures, and strategies, and identifying design features, display formats, etc. to correspond to and support the users' cognitive structures.

3. Empirical Results

This section presents a survey of empirical findings pertinent to the theoretical issues discussed in Section 2. The intent is not to provide a comprehensive review of the vast empirical findings documented in the literature but to address issues raised in previous sections and determine recommendations about the applicability of theoretical concepts. In addition, the empirical results reviewed will indicate important lessons learned that should be considered in the upgrade of the HALIFAX class.

3.1 Decision Making

A problem for the evaluation of analytic and intuitive theories of decision making is that both classes of theory have received support in numerous studies (e.g., Adelman, 1992; Kaempf et al., 1992; Leedom, Adelman, & Murphy, 1998; Newell, 1990). This precludes a simple distinction between these theoretical perspectives. The best that can be done is to identify the situations and conditions under which analytic theories serve as the better model of human decision making and the situations and conditions under which intuitive theories serve as the better model.

3.1.1 General Decision Making Processes in C2 Environments

A number of studies have sought to characterize human decision making in naval C2 environments. These studies provide a broad description of what tactical decision makers do. Serfaty et al. (1997), for example, conducted 20 US Assistant Sensor Weapons Controller (ASWC) officers in a low-fidelity simulation of anti-submarine warfare scenarios. The participants attempted to detect target submarines and estimate their position and heading. Serfaty et al. (1997) concluded that ASWCs were following a three-stage process of matching the situation to a schema (recognition), gathering information to elaborate the schema, and then recognizing a plan for action. Many studies (e.g., Cannon-Bowers & Bell, 1997; Kaempf et al., 1992; Klein et al., 1995) have also found that naval C2 decision makers focus on recognizing the situation.

Serfaty et al. (1997) also performed a correlational study with Army commanders and observed five general decision making steps:

1. Generate a schema and initial plan.
2. Ask the right questions (seek data).
3. Develop a rich mental model.
4. Visualize outcomes (simulate COAs).
5. Develop and select a COA.

In general, these steps are consistent with intuitive theories of decision making (e.g., Kaempf et al., 1992).[3] Other studies examining the constraints on decision making have suggested the reasons why human decision makers seem to perform intuitively. Hutchins (1997), for example, reports an experiment examining decision making in simulation scenarios of an

[3] These points can also be somewhat consistent with analytic decision making if multiple schemata and outcomes are considered.

AEGIS class cruiser. She observed that the volume of information in the scenarios (typically 11 to 18 contacts of interest at any given time) exceeded operators' memory and attentional capabilities. Participants had difficulty maintaining SA and had little cognitive capability to accomplish other tasks. Thus, it seems that under realistic conditions, naval C2 personnel will be unable to concurrently consider multiple hypotheses or COAs due to limits of memory and attention.

Intuitive theories have an advantage over analytic theories in explaining decision making in demanding situations because they involve less computation and make use of more automatic processes (i.e., recognition, cue association, etc.). Thus, it is not surprising that tactical decision makers seem to exhibit behaviour consistent with intuitive theories. Some studies, however, suggest that these findings are partially due to where researchers have been looking in the decision making process as much as human decision making strategies. Roth (1997) performed a study of decision making by nuclear power station crews. Although this domain is different from the naval tactical domain, it shares several important characteristics. In particular, it is a high-risk domain where crews may be called upon to make decisions under uncertainty and severe time constraints. Roth (1997) analyzed performance of crews in simulated emergencies and observed two kinds of activity. The first consisted of monitoring and developing SA in an attempt to recognize the situation. This is consistent with the activity observed in naval C2 teams. In addition, Roth (1997) observed procedural activity in which crews attempted to apply pre-planned responses. This is, in a sense, an intuitive strategy but it relies on analytic approaches in a prepatory phase. Thus, crews spent time before confronting the emergency in developing procedures to deal with emergency situations. The Roth (1997) study suggests that researchers should consider not just crisis situations when studying decision making.

Webb and McLean (1997) documented mission planning and preparation aboard the CPF, listing the critical steps taken to clarify missions and ensure the readiness of all personnel. OR teams aboard the CPF spend a great deal of time planning missions and anticipating potential situations that might arise. Mission planning and preparation is a distinct phase of operations, defined by the need to acquire information about expected threats. Much of this activity focuses on establishing plans to be communicated to subordinates. The plans prepare the ship and crew to meet anticipated threats by providing detailed analysis of enemy and friendly force capabilities, the political and civil situations, neutral forces, and parameters of the mission (ROE, political resolve, etc.). Success in generating plans depends on acquiring current and accurate information and so time is spent to ensure this.

In addition to mission planning, OR teams will have had extensive training prior to posting aboard a vessel. Planning and training are situations in which tactical decision makers have sufficient time to engage in analytic strategies to consider and solve potential problems.

3.1.2 Applicability of Analytic and Intuitive Theories of Decision Making

Recently, researchers have begun to consider not which theory describes human decision making but the conditions under which theories describe human decision making. Several studies have sought to determine the specific decision making strategies used by tactical decision makers. In general, these studies have considered the breakdown of strategies in general, seeking to determine the frequency with which intuitive and analytic strategies are used in tactical scenarios. Three studies, in particular, illustrate this work.

Pascual and Henderson (1997) videotaped and audiotaped naval C2 teams as they performed scenarios. They collected written outputs (notes, diagrams, maps, etc.) and self-reports of factors such as subjective workload. The video and audiotapes were then used in a modified Critical Incident Method (CIM) (Klein, Calderwood, & MacGregor, 1989, cited in Pascual & Henderson, 1997). Subject Matter Experts (SMEs) watched the videotapes of sessions and recorded critical incidents and decisions of interest.

The coding of this data indicated a problem with unambiguously attributing one particular model of decision making to observed behaviour. This is the result of both the ambiguity of the recorded behaviour and the overlap between models. Nevertheless, Pascual and Henderson (1997) were able to identify behaviours consistent with a number of theories, including:

- RPD (Klein, 1992, 1997).
- Image Theory (Beach, 1990).
- Schemata Model (Noble, 1986, cited in Pascual & Henderson, 1997).
- Story Model (Pennington & Hastie, 1986).
- Risk Assessment Model (MacCimmon & Wehrung, 1986; cited in Pascual & Henderson, 1997).
- Analytic Holistic Model (Rapoport & Wallestin, 1972).

RPD, Image Theory, and the Schemata Model are all intuitive theories, whereas the Risk Assessment Model and Analytic Holistic Model are analytic theories. In general, participants in Pascual and Henderson's (1997) study employed intuitive strategies 87% of the time and analytic strategies only 2% of the time (hybrid models 3%). This result is consistent with research in general that confirms intuitive models of decision making.

Kaempf et al. (1992) conducted a similar study to identify the decision making strategies employed by AEGIS class cruiser commanders while performing a threat assessment task. They found that 95% of situation assessment decisions were made by a recognition-based process. Only 5% of decisions indicated an analytic strategy. In addition, the most frequent diagnostic strategy was feature matching (88%), in which commanders made use of just a few clear features (e.g., bearing, response to warnings, etc.) to recognize the kind of target indicated. Commanders did use mental simulation in 11% of decisions, particularly the most difficult.

Finally, a study by Leedom et al. (1998) also indicates the general preference for intuitive strategies but also the use of some analytic strategies. Their study was performed with Army commanders during a four-day simulation of move to contact, defence, and attack missions. They conducted critical decision interviews with participants to identify key decisions and the strategies employed by participants.

Leedom et al. (1998) found evidence supporting three particular theories. The first was RPD (Klein, 1992, Klein et al., 1993). Participants frequently attempted to match the current situation to a template or memories of past experiences. In addition, they frequently engaged in story building to build a mental model of enemy maneuvers and options.

The second theory obtaining support was Image Theory (Beach, 1990). Image Theory involves the constant monitoring of the situation to determine whether intervention is required. Many of the decisions made by participants focused on this and on obtaining the

data needed to assess the situation. Participants tended to focus on one COA at a time and relied on their experience to generate and evaluate COAs.

The third theory supported was an analytic model proposed by Zsambok et al. (1992, cited in Leedom et al., 1998). This model involves the incremental shaping of the problem space to isolate the best solution path. In the context of the battlefield exercise, this entailed determining how to use resources (landmines, air support, etc.) to alter the tactical situation to achieve a state closer to the goal state.

The results of all these studies indicate that intuitive models, in general, provide a good description of what tactical decision makers do. Evidence supports a number of specific theories within the intuitive classification so it is probably not worthwhile to try to distinguish one particular intuitive theory as the best or most suitable. Rather, effort should be directed to identifying the processes and strategies used specifically by C2 teams aboard the CPF. Behaviour of decision makers in the field will be best described by taking theoretical concepts from a range of particular theories.

The results of the study also demonstrate some use of analytic processes. This is definitely less common than the use of intuitive strategies but does constitute a part of the tactical decision maker's approach. Thus, researchers should not ignore the use of analytic strategies when describing human decision making or designing decision support. The work of Roth (1997) and a study by Xiao, Milgram, and Doyle (1997) of anesthesiologists both suggest that analytic strategies might play an even bigger role when decision making is considered in a broader context.

Xiao et al. (1997) observed that anesthesiologists engage in a great deal of preparation before an operation. The goal of this preparation is to identify potential problems that could arise and to identify procedures to deal with those problems. Especially important was identifying a list of Points For Consideration (PFCs) that indicated special conditions, troublesome scenarios, obstacles to routine procedures, possible errors, and so on. This allowed the anesthesiologists to set up warning flags and prepared responses so that emergencies could be dealt with quickly and efficiently.

The strategy of anesthesiologists demonstrates an integration of analytic and intuitive decision making. During the preparation phase, they conducted a thorough review of the operation and could consider multiple solutions to anticipated problems. During the operation itself, they relied on recognition strategies that were facilitated by their preparation.

Of immediate relevance, Webb and McLean (1997) observed extensive preparation by OR crews aboard the CPF. The preparation phase before a mission was marked by the need to acquire information about expected threats and to prepare the ship and crew to meet those threats. Thus, the ability to generate COAs and evaluate them is crucial. Although Webb and McLean's (1997) study did not examine the cognitive strategies underlying planning, it seems that analytic strategies could enhance the OR team's ability to predict threats and determine the best possible response.

The Webb and McLean (1997) study provides insight into the broader decision making process of the OR team but further researcher should investigate the nature of these preparation activities and formulate a theoretical account. Virtually every study of C2

decision making deals with the crisis situation only and does not simulate the preparation that is done in mission planning.

3.1.3 Factors Affecting Decision Making

Consistent with this analysis, research has identified a number of factors that determine the strategies used by decision makers. A comprehensive review of this research is beyond the scope of this report. Instead, this section will summarize the relevant factors and discuss the situations in which decision makers might favor intuitive strategies and those in which decision makers might favor analytic strategies.

Perhaps the most important of these are factors related to the task. The nature of the situation can severely limit what a decision maker does. Among the task-related factors are (McMenamin, 1995; Flin, 1998; Pascual & Henderson, 1997; Klein, 1992):

- Available time.
- Decision timing (when decisions are required).
- Data quality/availability.
- Level of risk.

Of these, perhaps the most frequently noted factor is the time available to make a decision. Limits on the available time make it very difficult to employ an analytic strategy because the decision maker is unable to accomplish all the necessary computations. Similarly, decision timing can limit analytic strategies if several decisions must be made in rapid succession. As noted earlier (section 2.2.3.1), analytic strategies require relatively complete, high quality data to generate solutions. High levels of risk also require decision makers to respond quickly with a workable solution to reduce the risk.

Although, intuitive processes are favored for tasks that involved limited time to make decisions, rapid pacing of decisions, uncertainty or ambiguous data, and high levels of risk, analytic theories may be favoured for the strategic level of C2. The strategic level involves more planning activities and generally functions under less severe time constraints and with higher quality data. Even at the tactical level, planning and mission analysis can occur at a slower pace, making analytic processes more useful.

In addition to task factors, a number of individual factors affect the use of decision strategies. These include (McMenamin, 1995; Flin, 1998; Pascual & Henderson, 1997; Klein, 1992):

- Workload.
- Familiarity of the situation.
- Experience.

The first two factors, workload and familiarity, are actually joint factors of the situation and the individual. They are, however, expressed through the decision maker. The greater the workload, the less effort the decision maker is able to devote to conscious, deliberative reasoning processes. Thus, automatic processes, such as recognition and other memory-based strategies, are very useful when a person is forced to make a decision while doing other tasks. The familiarity of the situation, however, will determine the success of these strategies. Intuitive strategies tend to break down when the decision maker cannot recognize similar situations. Related to this factor is the experience of the decision maker. This determines the store of memories and schemata that can be used to categorize the current situation and retrieve a workable COA.

If the situation is unfamiliar and/or the decision maker lacks experience, he or she can still employ intuitive strategies such as story building but they will be less effective. Analytic strategies, by their nature, are general problem solving techniques and can be applied even in unfamiliar situations. In practice, however, people seem to apply formal problem solutions and logic more effectively when the situation is familiar (Cheng & Holyoak, 1985; Gigerenzer & Hug, 1992).

3.1.4 Expert Versus Novice Differences

The role of expertise in decision making deserves additional attention because it is crucial to intuitive strategies and because researchers have devoted a fair amount of research to it.

Some research has attempted to characterize the nature of expertise. Experts, of course, are individuals who have performed in the domain longer than novices and experts exhibit higher levels of performance. The crucial issue concerns the cognitive structures that experts have acquired through practice that lead to the better performance. Thus, researchers examining performance in domains ranging from chess (de Groot, 1965) to solving physics problems (e.g., Glaser & Chi, 1988, cited in Anderson, 1995) have attempted to identify the specific ways in which experts differ from novices.

Elements of Expertise
• Experts perceive great numbers of interpretable patterns in their subject matter, indicating excellent organization of knowledge
• Experts perform faster and with fewer errors than novices
• Expert possess superior short-term memory (STM) and long-term memory (LTM)
• Experts perceive and represent problems in domain at a deeper, more conceptual, level than novices
• Experts quantitatively analyze problems to represent them mentally and specify situations and constraints
• Experts possess substantial self-monitoring skills when detecting errors, checking solutions
• Experts spend proportionately more time than novices constructing problem representations
• Experts possess larger sets of schemata than novices
• Expert knowledge is organized to greater degree so it is more accessible, functional, efficient to use
• Experts' problem perception is schema-driven, whereas novices' problem perception is based on general search strategies
• Experts exhibit more context-dependent performance than novices; novices possess context-free features and facts and rules for behaving

Table 3.1 – Elements of Expertise (from Federico, 1995)

Some of the essential elements of expertise (see Federico, 1995) are listed in Table 3.1. Many of these elements are basically descriptive. Experts perform faster and that is a feature by which we recognize expertise. Other elements are difficult to interpret. Experts possess better short-term and long-term memory skills but one can ask whether this indicates that memory improves with practice or that superior memory skills allow individuals to become

experts. Still other elements indicate a degree of refining or improvement in certain cognitive skills. Thus, experts exhibit better organization of knowledge so that they are able to mentally group related concepts and create integrated representations of problems.

Perhaps the major element of expertise is the acquisition of knowledge. Simply by practice, the expert encounters many situations and attempts many solutions to problems. In the process, this information is recorded in memory and organized. Given that decision makers often rely on intuitive strategies, this store of information is extremely valuable.

The organization of knowledge, of course, is another key factor. In particular, expertise is developed when a person relates domain knowledge in terms of conceptual rather than surface elements. Experts recognize more meaningful aspects of problems and assign greater weight to the conceptual structure of the situation than novices (Federico, 1995). That is, experts look for features related to functions, goals, intent, and so on, and use these features to identify the problem as well as retrieve appropriate solutions. Novices tend to look for the most salient, observable features that indicate some surface similarity to previous experience. The greater use of conceptual organization by experts allows them to identify solutions that address the meaningful issues of the problem.

Kirchenbaum (1992) performed a study that demonstrates this difference between expert and novice problem solving. She observed participants in a submarine warfare scenario and classified the strategies by which participants sought information to determine a COA (attack, surveillance, or retreat). The strategies were:

- Quantity: survey as large a set of data as possible.
- Ease: examine data according to the ease of access.
- Relook: go back to previously examined data.
- Class: search for conceptual structure in processed data rather than surface structure in raw data.
- History: look for information pertaining to changes in the situation.
- Set: identify chunks of related information.

Kirchenbaum (1992) found that experts tended to use the history, set, and class strategies more than novices. Experts also used a great deal of raw data but used less data overall than novices. Novices exhibited greater use of the ease strategy. Overall, experts were better able to look for and extract the meaningful patterns in the data, which led to better performance in the scenario task.

Differences in memory organization are especially pertinent to naval C2. Given that many situations will favor, if not require, intuitive decision making, tactical decision makers will rely on their memories a great deal. In general, experts' well-developed conceptual understanding of the problem domain aids recognition of critical cues and classification.

It is important to note that experts and novices do not necessarily differ in the *amount* of conceptual or structural data they use when solving problems (Federico, 1995). Often, experts and novices sample the same data. Differences in their performance result from how experts and novices use data. Novices tend to perform in a *context-free manner*. That is, they fail to connect the current situation to previous experience. Experts may not process the data at hand to any greater degree than novices but they are able to link it quickly to their domain knowledge.

3.1.5 Decision Making Styles

Another factor that could potentially affect tactical decision making is the decision making style of the commander. Individuals, especially those in leadership roles, can exercise a great deal of personal choice over how they accomplish tasks. Flin (1998) has identified four main styles used by commanders.

Creative. This strategy is based on knowledge-based reasoning. The commander employs intuitive processes such as feature matching and recognition but attempts to synthesize several ideas to generate a novel solution. This approach requires a great deal of WM and attentional resources and is more likely to be used during planning than in a crisis situation.

Analytic Option Comparison. This style emphasizes analytic processes of problem space definition and comparison of COAs.

Procedural. This style relies on the implementation of standard operating procedures with an emphasis on situation assessment and determining the appropriate procedure to apply.

Intuitive. This strategy emphasizes the rapid recognition of the situation based on experience and the retrieval of a satisfactory COA.

No commander employs strictly one style; commanders use all four to varying extents based on the situation and task factors discussed above. Nevertheless, commanders are likely to exhibit some personal preference for certain approaches. A potential problem occurs when personal preference conflicts with the demands of the situation.

3.1.6 Implications of Research on Decision Making

Several important implications arise from the empirical work on decision making. First, performance in tactical situations is severely constrained by the limits of human memory and attention (e.g., Hutchins, 1997). In particular, people have a small short-term memory capacity that restricts how much information they can process at any one time. The flip side of this problem is that tactical situations confront the decision maker with vast quantities of data. Consequently, researchers have proposed support systems to reduce memory and attention loads. Under the Tactical Decision Making Under Stress (TADMUS) project (e.g., Hutchins, 1997; Kelly et al., 1996; Morrison et al., in press), for example, decision support modules have been developed to graphically present historical, or time-related, data about tracks, and priority lists and alerts to give crew members immediate perceptual access to data. This reduces the need for the crew to remember relevant data. Note, however, there is a possibility that increasing the amount of data on a display can impair attention.

Another major finding has been that tactical decision makers tend to rely on intuitive processes to a great extent. Thus, decision makers employ recognition, feature mapping, story building and so on much more often than analytic processes such as multiple attribute comparison or problem space shaping. This indicates a need to support recognitional processes by identifying and highlighting salient cues and augmenting human memory with databases (e.g., Leedom et al., 1998). One drawback of intuitive processes is that they are subject to memory limitations. Thus, decision makers can rarely consider more than one hypothesis at a time (Klein, 1992; Klein et al., 1995). Thus, there is also a need to help decision makers work around this limitation to provide the capability to simultaneously evaluate multiple hypotheses.

Empirical studies and, to an extent, theoretical analyses have tended to ignore the planning stages of C2. Thus, our understanding of how C2 crews prepare for missions is underdeveloped. In particular, this seems to be an area where analytic theories of decision making could be most useful. Ultimately, to fully support C2 crews in all activities, it may be necessary to combine intuitive and analytic approaches. Analytic-based support tools could assist crews in planning and developing prepared responses to anticipated situations. These tools would also help crews develop sets of cues by which to recognize situations and retrieve the pre-planned responses. Intuitive-based DSSs would assist crews in situation assessment, feature-mapping and recognition, and retrieval of responses.

3.1.7 Recommendations

- Provide decision makers with support to their limited WM and attention but ensure that support to WM does not increase attentional demands
- Provide support to intuitive decision making strategies (e.g., in the form of intuitive-based DSSs) for rapid response.
- Conduct further research on the role of planning in tactical decision making with an aim to designing DSS concepts to support mission planning and improving the integration of planning and response in crisis situations.

3.2 Decision Support Systems

A great deal of work has been done to create working DSSs, including DSSs for naval C2 and tactical decision making. This section will review some general findings regarding the development of DSSs and discuss in more detail several specific DSSs relevant to naval C2.

3.2.1 Approaches

A DSS is often defined as a computer-based system that assists human decision makers to deal with ill-defined problems and select a COA (e.g., Sprague & Carlson, 1982, cited in Hart, 1988). It seems that decision support could be defined in a much broader sense to include the non-computer-based tools (e.g., paper and pencil, whiteboard, manuals, etc.) available to a decision maker, as well as human resources such as assistants or other teammates. Research in the area, however, has focused on computer aids, probably because of the great potential of computers to perform decision-related tasks quickly, efficiently, and accurately.[4]

DSSs typically consist of three main functional components, a database, a model-base, and the decision maker (Sprague & Watson, 1983, cited in Manning, 1991). The database contains all the domain-related knowledge that the computer system will need to process data and either make recommendations or present processed data to the decision maker. The model-base consists of the formal methods by which the system manipulates data. The model reflects designers best understanding of the domain and the needs of the decision maker. The decision maker is an integral component of the system itself and retains supervisory and veto control. In some approaches to DSS design, the decision maker is the locus of decision making activities, performing most of the relevant computations. Even in approaches that

[4] Note, most work on the development of DSSs has focused on control rather than command (see Pigeau & McCann, 1995).

employ a great deal of computer computation, the human decision maker ultimately determines how to apply the system's output.

There are three broad approaches to DSS design (Hair & Pickslay, 1992).

Expert System (ES) DSSs. The first approach is to design a computer system that employs formal analytic methods to solve problems. This is the *expert system* approach in which the system plays the role of a decision maker (Hart, 1988; Hutchins et al., 1996), although the way the computer system solves problems will likely not be the way human decision makers solve the problems. Instead, the computer's speed and computational power allow it to use highly analytic methods to analyze data and select a COA. Because of the expert system's use of analytic methods, it will generate an optimal solution (given the completeness and accuracy of data and appropriate algorithms for the circumstances). The human decision maker's role is strictly supervisory. The human retains final authority and can accept or veto the system's recommendation. Expert systems can be described as relatively restrictive DSSs (Glover, Prawitt & Spilken, 1997) in that they limit the user's decision making processes. They also provide little or no decision guidance to help the user choose a COA from a set of options. Although computers have great computational power (and can expected to increase in power in the near future), analytic methods can still be time consuming, especially if the system acquires large amounts of data. Furthermore, these systems tend to be highly complex so that the user will probably not understand how the system works or achieves its recommendation. This can be a great drawback for the human decision maker who will have little basis for determining whether the system's recommendation is a viable one or not.

Interface DSSs. A second approach is to not use the computer as a reasoning tool but rather strictly as an interface between the human decision maker and data (Hair & Pickslay, 1992). In this type of system, the computer processes information to derive the best format to suit the human decision maker's needs. Its role is to identify relevant data attributes and select an appropriate model to summarize and present the information needed by the user (Manning, 1991). Thus, the DSS will display variables in graphic form to highlight important relationships, present alerts and other warnings, calculate higher order summary statistics, and so on (Hart, 1988). The goal is to create *cognitive fit* between the way information is presented and the strategies employed by the user (Section 3.6). In other words, the DSS processes data to the form that is immediately usable according to the intuitive and analytic processes of decision making. This sort of system involves relatively little computation and maintains speed. However, this sort of system does not take advantage of the full capabilities of the computer. It also places all of the computational responsibility on the human user, who will have limited cognitive capacity.

Intuitive DSSs. A third approach is to design a system that makes decisions the way a human does (Hair & Pickslay, 1992). This is similar to an expert system except that an intuitive DSS is designed to walk the human user through his or her own decision making rather than to replace human computation. It helps the user by performing highly computational tasks and presenting data in its most usable format (Manning, 1991).

In this approach, the computer employs more intuitive, informal strategies (e.g., Kolodner, 1991, cited in Hair & Pickslay, 1992). Thus, the system is understandable to the user. The computer is, however, not subject to the biases that humans are and so can perform these strategies more accurately and certainly faster than a human decision maker. Because the

system uses the same kind of reasoning processes, the human decision maker will find it easier to incorporate the system's computations within his or her decision making process. The user will also be better able to evaluate the trustworthiness of the system's outputs because its methods are more readily understood. This kind of DSS is less restrictive, in general, than an ES, allowing greater user input and offering more decision guidance (Glover et al., 1997).

Communications support is another aspect of DSSs especially pertinent to naval C2. Increasingly, researchers have been investigating team-level decision support (e.g., O'Neill, 1996). This form of support can consist of anything that improves the flow of information between team members, such as computer-based systems to pass text or graphics objects or allow real-time interactions between physically disparate people.

3.2.2 Characteristics of DSSs

Based on the above analysis, the two major functions of DSSs are to give the user needed information in the needed format and to take over computational tasks for the user (Morrison et al., in press). More specifically, the first function is to help the user integrate and evaluate large amounts of data and to develop a coherent mental picture of the situation (Urban, 1990). The second function is necessary to deal with limits of human cognitive processing capacity (Hutchins, 1996). By performing computations for the user, the DSS reduces the human decision maker's workload and allows him or her to perform other tasks.

Design features of DSSs are arranged to accomplish these functions. Sprague and Carlson, 1982, cited in Hart, 1988) list four main design concerns related to these functions:

- Representations: how the situation is represented through the data.
- Operations: capabilities of the system that the user can employ.
- Memory aids: methods to remind the user of important information.
- Control mechanisms: option menus and user-system dialog that allows the user to direct the DSS.

Thus, design of a DSS must include specification of:

- An interface that facilitates situation understanding.
- Tools that improve decision making operations.
- Database and informational resources that the user can consult.
- Control options that allow the user to manipulate how the system performs.

Neglecting any of these areas will limit the usefulness of a DSS.

In addition, there are several task-related functions of decision makers that also need support. Some DSSs are specifically designed to enhance communication and team interaction (e.g., O'Neill, 1996). This form of support can consist of anything that improves the flow of information between team members, such as computer systems to pass text or graphics objects or allow real-time interactions between physically disparate people (see Section 3.2.8). The more efficiently team members can pass needed information, the better will be their representation of the tactical picture (Webb & McLean, 1997). Likewise, individual SA is crucial (Hutchins, 1996). Design of interface and tools can help the user monitor events and create a mental model of the situation.

Control options are important because the user must be able to adapt to changing situations. C2 teams cannot pre-define and plan for all situations, especially in the modern littoral environment (O'Neill, 1996). Thus, tactical decision makers will be confronted with changing and often unfamiliar situations that require changes in how data is processed and used in decision making.

3.2.3 Support of Intuitive Decision Making Strategies

If DSSs are to emulate the way humans make decisions or provide cognitive fit, they must support intuitive strategies. The primary aspects of intuitive strategies are (Hutchins, 1996; Kaempf et al., 1996):

- SA.
- Feature matching.
- Recognition.
- Selecting a COA.
- Story generation or mental simulation.

The goal is to build a complete picture of the situation, categorize the data, match the current situation to a schema or template or checklist, choose an associated COA, and test the validity of the COA by simulating its consequences (i.e., by visualizing the "story" ahead, or story-building). Thus, intuitive theories indicate that these are the processes in need of support.

3.2.4 Support of Analytic Decision Making Strategies

Tactical decision makers use analytic strategies to a lesser extent but still require some support in this area. Analytic processes include:

- Problem space definition.
- Multiple option generation.
- Option comparison.
- Optimization criteria.

Decision makers need support in all these areas but multiple option generation and comparison are probably the most difficult for human decision makers. Thus, a DSS must be based on a good model of the domain so that it can perform much of this work and suggest reasonable alternatives (Gabrielson, 1995).

One way to expand a DSS's model of the domain is through user input. Adelman, Cohen, Bresnich, Cinnes, and Lashey (1993) describe an expert system to assist US Army air defense officers identify incoming aircraft and engage foes. A feature of this system is it rule-generation capability by which the operator defines rules for classifying contacts. The system helps the operator determine relevant criteria and implement the rules that can then be evaluated automatically by the system. Thus, the system enhanced the operator's understanding of the problem domain and performed the computationally challenging task of evaluating the rules. Rule-generation improved the performance of air defence officers in identifying and engaging foes in simulated encounters (Adelman et al., 1993).

Analytic support may be particularly useful for inexperienced personnel who have not developed familiarity with tactical situations (Hutchins, 1996). Novices operate more on the basis of a declarative, formal understanding of the domain than experts (Anderson, 1995) and need support to help translate the formal rules into action.

3.2.5 Practical Issues

To be effective, a DSS must not only perform well but be *perceived* to perform well by its users. If user acceptance is low, the use of a DSS might have to be mandated. This means that extensive user reviews must be conducted during rapid prototyping (Section 5.1.1.3) of systems to ensure that the DSS will be accepted. Hart (1988), for example, interviewed 14 experienced Tactical Action Officers (TAOs) serving aboard AEGIS class cruisers to identify user concerns for DSSs. The TAOs were almost unanimous in their belief that a DSS would have to be fast and accurate to be useful. In other respects, however, there was greater disagreement. Some TAOs preferred that the system provide a fast recommendation rather than pursue multiple options but others preferred that the system seek the most accurate solution. Most preferred that the system not limit the number of data tracks that could be monitored but they disagreed about how to identify tracks. Gaining the acceptance of all or most users may be difficult but an accepted system likely will be used more efficiently.

Another issue to consider is that decision makers will work long shifts during which they must pay close attention to the unfolding situation. This sort of sustained attention can be difficult and lead to loss of vigilance. Hollenbeck, Ilgen, and Weissbein (1996) observed that tactical crews were especially susceptible to lapses of attention when working with systems that performed many tasks for the user. They advocated the somewhat counter-intuitive suggestion that systems require users to manually chart large amounts of data. The purpose of this is to force the user to cognitively process data and maintain attention to the task.

An important aspect of a DSS is how it presents information to the user and the extent to which the format creates cognitive fit between the displayed information and the user's mental model of the task. Consequently, HCI issues are important to producing an effective DSS (see Miller et al., 1992). One particular technique that enhances comprehension in general is the use of a graphical format. Graphics can simplify information and improve its legibility. Moreover, it is often easier to see relationships between variables when they are plotted in a visuo-spatial form (e.g., Jarvenpaa, 1989). Hutchins (1996) notes two main benefits of graphical presentation to the user of the DSS. First, graphical formats reduce the amount of mental computation because some of that work has been done for the user. The graph allows users to use less demanding perceptual processes to see important data and relationships rather than infer or compute them from text or tabular forms. Second, it can allow the user to spend less time looking for information if the display is less cluttered and better organized.

3.2.6 Potential Problems of DSSs

DSSs can have negative effects on the user's behaviour. Some of these have been mentioned, notably that the user may become complacent and simply accept the recommendations of the DSS without considering the problem. A highly automatized DSS has the potential to encourage users to minimize their effort, which in turn can reduce attention and vigilance (Glover et al., 1997). DSSs can also induce users to treat their tasks mechanistically. That is, users begin to simply follow set procedures without thinking conceptually about the task. This sort of performance undermines supervisory control of the system and fails to take advantage of expertise. In fact, it undermines the development of expertise as cognitive involvement is necessary to learn the large amounts of associative and procedural knowledge underlying expertise (see Anderson, 1995).

Another problem with over-reliance on the DSS is that the DSS may not generate accurate or useful recommendations. DSSs have the advantages of computational power and speed but are limited by the models and procedures built by designers. The human operator will be better able to adapt to the specifics of the situation.

3.2.7 Examples

This section reviews several specific DSSs for naval C2 that illustrate the issues just discussed. The DSSs reviewed are based, in general, on intuitive theories of decision making and emphasize the role of the human decision maker rather than the generation of recommended COAs. These examples are not necessarily representative of the range of DSSs that have been developed over the years (see Wagner, 1989, for a review of some earlier naval DSSs) but do convey some important developments in decision support. Note that these DSSs were developed under the US Navy and represent a technological and control oriented approach to C2. DSS projects of other nations could provide a useful counterpoint to the DSSs described here.

3.2.7.1 Situation Assessment By Explanation-based Reasoning (SABER)

SABER was developed under the TADMUS project and is, in fact, incorporated in the TADMUS DSS discussed below. It is an intuitive DSS that explicitly attempts to emulate human decision making processes so that the decision support is understandable by the user and can guide the decision maker away from biases and errors (Hair & Pickslay, 1992).

The role of SABER is to help the operator (SABER was developed for use at multiple positions) detect and identify sensor contacts. It employs explanation-based reasoning processes, which have been proposed in intuitive theories of decision making (Hair & Pickslay, 1993). The goal of SABER is to generate explanations for sensor tracks that best account for the available data. It provides multiple explanations to the user to help him or her evaluate possible identifications of tracks. Explanations take the form of simple causal models that indicate what outcome is expected given the data available. Figure 3.1 shows an example of the examples display for SABER.

Ordered conclusions: hostile threat, commercial, hostile routine, friendly

Confidence: possible

EXPLANATIONS

1. hostile threat	**2. commercial**
DIRECT SUPPORT: no IFF, no radar	DIRECT SUPPORT: air corridor C
ASSUMPTIONS: About air path C: hostile aircraft is imitating C hostile is off course	ASSUMPTIONS: About no IFF: iff malfunctioning About no radar: radar malfunctioning

3. hostile routine	**4. friendly**
DIRECT SUPPORT: none	DIRECT SUPPORT: none
ASSUMPTIONS: About air path C: hostile aircraft is imitating C About no IFF: hostile only testing our response About no radar: hostile only testing our response	ASSUMPTIONS: About air path C: friendly is off course About no IFF: iff malfunctioning About no radar: radar malfunctioning

Figure 3.1 – SABER Explanations for Observed Sensor Data (from Hair & Pickslay, 1992)

A basic explanation indicates that a certain conclusion is warranted because certain data is present (or that a certain conclusion is not warranted because certain data is absent). When several kinds of data are present, individual explanations can be combined into larger explanatory structures. Thus, SABER combines separate pieces of data and determines which conclusions can satisfy all available data. It also provides a confidence level (confirmed, probable, or possible) to indicate the strength of the causal links between data and conclusions.

SABER also provides an evaluation of the explanations based on simplicity, completeness, and data significance. For the last criteria, SABER evaluates the weight that should be assigned to each piece of data when developing its explanation. This requires a good model of the task to determine ahead of time what kinds of data are

expected to occur with each kind of sensor contact (Hair & Pickslay, 1992). Prior to deployment of SABER, designers must specify the set of possible conclusions and the set of possible kinds of input data. The latter set is easier to specify because it is based on the sensor equipment of the ship.

The SABER database is user-modifiable (Hair & Pickslay, 1992). The user can enter new explanatory structures, indicating which conclusions are linked to certain data, and change the weights used to evaluate explanations, thus altering how strongly certain data indicate certain conclusions. In this way, SABER can accommodate the expertise of the user in the field and the user's perception of changing conditions.

SABER also incorporates a *critic* function (Hair & Pickslay, 1993). This provides the user with criticism of his or her on-going decision process to improve the quality of the solution that is adopted. The SABER critic focuses the user's attention on salient pieces of data (as judged by the SABER database) to ensure that the user does not overlook data due to cognitive biases. It can also highlight alternative explanations if the user indicates a preference for a particular explanation. The critic function requires an extensive model of the user so that the system can predict the user's needs and highlight only information actually needed by the user (Hair & Pickslay, 1993).

3.2.7.2 TADMUS DSS

The TADMUS project was undertaken after two tragedies struck the US Navy. The USS *Stark* was severely damaged by an attack and 27 US naval personnel lost their lives when the commander decided not to engage an inbound aircraft that was judged not to be a threat (Hutchins, Morrison, & Kelly, 1996). In contrast, the commander of the USS *Vincennes* made the decision to engage an inbound aircraft that was judged to be a threat, an aircraft that turned out to be a commercial airliner, resulting in the deaths of all personnel aboard the aircraft (Hutchins et al., 1996). The purpose of the TADMUS project is to conduct research in the areas of human factors and decision making to find applications that can help avoid these types of events in the future. The TADMUS DSS was designed to support the decision making of command-level (CO and TAO) decision makers aboard the AEGIS class cruiser (Hutchins, 1996). The overall goal of the TADMUS DSS is to improve SA and decision making by helping personnel identify critical contacts earlier and more accurately and to determine and initiate appropriate responses at the right time (Morrison et al., in press).

The TADMUS DSS is based on the principle that modern warfare places a premium on rapid identification and COA selection (Hutchins, 1996). Kaempf et al. (1993), for example, found that in an anti-air mission simulation, 103 of 183 decisions involved situation assessment. Thus, a primary function of the DSS is to help the operator focus on relevant data. It also reduces the operator's mental workload by performing some information and data processing. The processed data allows the operator to *see* important variables and relationships rather than search for or infer them from tabular or text data (e.g., Hutchins, 1997).

The general design principles of the TADMUS DSS are to (Hutchins, 1996):

- Help the user anticipate future events by portraying contact history.
- Help the user identify stored patterns of events.
- Display planned responses for anticipated events to facilitate COA selection.

In other words, the DSS was designed to provide indicators of relations between data features, structural similarity between sequences of features, and evidence of relationships.

The designers of the TADMUS DSS applied two intuitive models of decision making, namely feature matching and story generation (Hutchins, 1996; Hutchins, Morrison & Kelly, 1996). The DSS helps the operator categorize incoming data by processing data and presenting it in a form that highlights important features and relations. The operator uses the categorized data to match the current situation to a template based on experience. The DSS also helps the operator to construct causal explanations through the SABER window, which contains the SABER DSS described above.

The DSS is organized into six modules that carry out these functions (see Hutchins, 1996; Rummel, 1995).

Track Summary. This window presents contact parameter information.

Track History. This window depicts the speed, altitude, course, and range of a single contact on a Two Dimensional (2D) display with contact and ownship weapon ranges indicated. These parameters are plotted against time so that changes in any of the parameters becomes apparent immediately. This facilitates use of a recognition strategy because the operator does not have to remember previous values to compare past states to the current state.

Response Manager. This window assists the operator in using preplanned responses by graphically depicting preplanned responses that need to be taken given the current data. In particular, it indicates the earliest and latest points at which specific actions (e.g., deploy chaff, fire missile) can be taken.

Basis for Assessment. This window contains SABER and uses explanation-based reasoning to generate possible accounts for the current data. It presents hypotheses regarding the contact and the assumptions for accepting each hypothesis.

Comparison to Normal Values. This window provides a quick comparison of features of a contact to features of examples of specific types of contacts to aid feature matching and recognition of contacts. The graphical display format aids quick reference to a knowledge base for identification.

Track Priority List and Alerts. This window focuses on several contacts at once and helps the operator monitor multiple contacts. It prioritizes contacts to indicate the most important to monitor and act upon. It also indicates the identification, probable intent, and actions that can be taken (with time envelopes in which each action can be taken).

The modules are arranged in order of information complexity from the top of the display to the bottom. At the top are the track summary, track history, and response manager windows, which all deal with analyzing and identifying a single contact. The basis for assessment and comparison to normal values provide more detailed analysis of a particular contact and are located in the middle of the display. The track priority list and alerts module presents information on all contacts requiring attention or action and is located at the bottom of the display.

The TADMUS DSS seems to be a successful system to support tactical decision making. Morrison et al. (in press) describe an experiment that assessed the impact of the DSS on the performance of COs and TAOs relative to the current system aboard the

AEGIS class cruiser. Experienced participants engaged in anti-air warfare scenarios in the DEFFT simulator. Participants performed two scenarios with the TADMUS DSS and two scenarios without it. Tactical actions, display usage, control inputs, and voice communications were recorded, as well as subjective assessments. Morrison et al. (in press) found that participants identified critical contacts earlier and more accurately when using the TADMUS DSS, except late in a scenario.[5] Participants also took more defensive actions required by the ROEs when using the TADMUS DSS. Furthermore, they also took those defensive actions earlier and took offensive actions later (thus avoiding escalation). Participants asked fewer questions to clarify previously reported track data and required less overall communication when using the DSS. Overall, participants rated the usability of the DSS as quite high. The results of this study indicate that the TADMUS DSS improves SA and contact identification and improves the effectiveness and timeliness of responses.

The TADMUS DSS is not a finished product and is undergoing development. Rummel (1995) reports a user evaluation of the DSS designed to assess user preferences for feature options. The evaluation revealed a number of user preferences, listed in Table 3.2, that may be relevant to the development of further DSSs.

3.2.7.3 Generic C3I Workstation

Anderson (1990) proposed a generic C3I (command, control, communication, and intelligence) workstation for the Composite Warfare Commander (CWC). This workstation is designed for installation in a wide variety of platforms and addresses decision support in:

- Information management.
- Platform management.
- Resource management.
- Tactical management.

Like the TADMUS DSS, the generic workstation has a modular design that permits rapid upgrades and modifications by designers. It also makes heavy use of graphical presentation to process data.

An important feature of the generic workstation is that it supports communications both within the OR team but also between the CWC and other platforms. The goal is to provide a common tactical picture and connect members across naval and land units (Anderson, 1990). It accomplishes this goal by providing the capability to process tactical data from many interfaces in real time. Thus, this workstation can help the CWC build a shared mental model with members of his or her own OR team and shared mental models with corresponding CWCs (or equivalent) on other vessels and land bases. This is potentially an important feature for the ORO aboard the CPF. The ORO communicates not just within the OR but with other OROs aboard other vessels, participating in the coordination of task groups.

[5] A possible explanation for the exception is that contact tracks may have become obvious after that amount of time so the DSS was not needed.

User Preferences for TADMUS DSS Options
Track Profile
Track priority list displays 3 most likely hypotheses, so window may not be useful
Some users may have no confidence in this window because the display would depend on a track's ID; if the ID is wrong, the window is wrong
Track History
Possible difficulties in using the display due to poorly defined platform attach profiles
Difficult to interpret the range display
Use a range scale rather than time scale because users convert ranges and speeds to time but not vice-versa
Alerts Window
Information should not be lost
Don't allow superseding alerts
Provide retrieval functions
ID is crucial and should be given its own window
Include warning status and weapons information
May want to order by track or threat level
Basis for Assessment
System's suggestions must be verifiable
Perhaps filter information that is common across hypotheses
Should contain information about weapons load, ID, and IFF
Users like unfiltered evidence lists
Response Manager
Should contain a readable list including a feedback function
One-dimensional response arrangement sufficient
Users liked one-dimensional arrangement but some preferred a time scale whereas others preferred a range scale
Feedback function is important
Track Priority List
Support least possible interaction
Priority lists are too slow and/or never update
Should focus on geoplot
Without range, it is impossible to automate the prioritization process
Bearing should be available automatically
ROE Support
ROE must be visible at any time, but not as a checklist

Table 3.2 – User Preferences for TADMUS DSS Features (from Rummel, 1995)

The generic workstation is based on six modules.

Communications interface. This module is responsible for transmission and acceptance of messages. It provides processing of messages to reduce the workload of the operator. In particular, the module processes messages to determine whether they a) contain track information, b) should be forwarded to others, or c) contain orders, actions, or other information that should be brought to the operator's attention.

Sensor interface. This module processes and interprets sensor data to provide an enhanced tactical picture.

Track database manager. This module is responsible for refreshing and updating all tracks. It permits rapid access to information fields about each track.

Track controller. This module automates the process of verifying track data integrity. The operator inputs constraints and track information. The module supports queries and information display requests to manage track data and handles the tasks of generating new tracks, deleting old tracks, and editing information in existing tracks.

Tactical command display. This module assimilates track and message data to facilitate tactical picture building. It also provides templates for displaying and editing communications messages.

Weapons systems interface. This module monitors weapon systems status.

Overall, the generic workstation is similar to the TADMUS DSS, taking the same approach, but contains greater communications functions.

3.2.8 Other DSS Concepts

In addition to these relatively well-developed DSSs, there are a number of DSS concepts or highly experimental systems that could be useful to the upgrade of the HALIFAX class.

3.2.8.1 Virtual Reality Command Centres

Recent technological advances provide the opportunity to create computerized environments to facilitate visualization of the battlespace, communication, and decision making. Both Dockery and Hill (1997) and Gardner et al. (1996) have proposed the use of some version of a Virtual Command Centre (VCC). A VCC is a virtual environment in which multiple participants at remote locations can interact in a shared Three Dimensional (3D) world. This technology can bring together individuals in physically separate command centers to improve the coordination of forces and contribute to a unified battle picture (Gardner et al., 1996). This technology can also be useful within the OR of a single ship where team members are spread throughout the OR and interactions are limited due to the layout of workstations.

The VCC proposed by Dockery and Hill (1997) would be a fully immersive environment, requiring the use of Head-Mounted Displays (HMDs). Within the virtual environment, the operator would have laid out all workstation and DSS functions. This aspect of VR could provide enhanced displays because the workspace can be arranged in three dimension. Furthermore, data can be presented three-dimensionally, increasing the sophistication of graphic displays. Interactions between individuals are supported by allowing individuals to enter the workspace of another person and see the same displays.

The ability to share workspaces could dramatically improve coordination of information for OR personnel aboard the HALIFAX class. In particular, it could help the ORO interact with members of the OR team aboard the ship but also other OROs aboard other ships. Virtual Reality (VR) technology could provide the ORO with two sophisticated communications channels so that the ORO could conveniently enter an ownship virtual environment or an abstract ORO workspace defined by the community of OROs in a taskforce.

VR poses several potential problems that could limit its applicability to naval C2. One problem concerns visual factors such as depth and slope perception and other sensual factors, such as perception of contact. Some research suggests that screen displays and VR impair depth perception. This is not surprising because current HMDs eliminate binocular depth cues, which are generally the most powerful for perceiving depth. Monocular or pictorial depth cues can be used to convey depth but it is unclear whether displays use enough of these cues. Some pictorial cues, such as converging lines and relative size, are easy to see in video displays, whereas others, such as texture gradients, may be difficult to see. Another common problem with VR is motion sickness. The main factors that affect this are the refresh rate of the HMD and the synthesis of visual motion with physical motion of the operator. People expect visual motion to be accompanied by physical motion and it may be necessary to simulate motion to avoid sickness and enhance the feeling of immersion.

3.2.8.2 Total Defence and Information System (TODAINFO)

TODAINFO is another concept to facilitate a unified tactical picture, developed as a Swedish Defence project. It has been developed in the context of joint forces C2 rather than naval C2 but the concept could be translated to the OR of the HALIFAX class. The main concept of TODAINFO consists of a 3D display of the tactical situation, which looks like a table-top aquarium, around which several participants can be seated (Sundin, 1996). An information technology (IT) operator operates the displays and manages communications.

TODAINFO was designed to receive data from local sensors, European surveillance satellites, and other TODAINFO terminals and creates a synthesized picture. Thus, the display consists of highly processed data suitable for high-level command. The situation is presented in terms of terrain, positions of units, weather, and tactical information (movements, etc.).

This kind of system could be valuable in implementing a "round-table" decision making environment. TODAINFO presents a comprehensive picture that multiple participants can view simultaneously for discussion and planning. It is unclear, however, whether this arrangement promotes the fast response times necessary for many C2 functions.

3.2.9 Recommendations

- Promote R&D to develop a DSS approach for the CPF that defines DSS requirements for representation, operations, memory aids, and control mechanisms, and determines which intuitive and analytic decision making strategies need to be supported at which stages of operations.

- Incorporate the practical issues of acceptability, sustained vigilance, and graphical format is the DSS design process.
- Employ current DSSs (such as the TADMUS DSS) as prototypes to help specify the requirements for a CPF DSS.

3.3 Situation Awareness

Relatively little empirical research has directly examined SA. Instead, researchers typically treat SA as a factor that affects decision making. This has limited insight into SA as a process in its own right, especially compared to the extensive empirical work on decision making. This section discusses two broad aspects of SA but more research is needed in this area.

3.3.1 Errors of SA

Endsley's (1995a, 1997) theory of SA provides a framework for categorizing common failures of SA (see Endsley, 1995a; Jones & Endsley, 1996). The loss of SA can reflect a number of different conditions and factors. Identifying the specific conditions and factors, however, provides design guidelines to help people maintain SA.

Level 1 errors. Errors at this level consist of failures to correctly perceive the situation, which can result from a number of factors:

- Inability to detect or discriminate data.
- Failure to monitor or observe the situation.
- Overabundance of information (exceed cognitive capacity).
- Inadequate data sampling strategy.
- Misperception of data.

Level 2 errors. Errors at this level consist of failures to comprehend the meaning of the situation. These errors indicate the absence of a complete mental model or the use of an incorrect mental model. Another component is the failure to recognize that one's mental model is inappropriate; i.e. one doesn't detect the mismatch between predictions or implications of the model to the current data.

Level 3 errors. Errors at these levels consist of failures to project the situation into the future. These errors can result from an inappropriate mental model or a lack of understanding of the dynamics of the situation. Also, limits of cognitive capacities can prevent one from mentally simulating future events.

Each level of errors indicates particular needs for SA support. Level 1 errors can be addressed by aids that enhance attention and WM, whereas errors at Level 2 can be addressed by aids that organize data and enhance recognition of situations. Level 3 errors seem to require support in these areas as well as tools that can simulate events.

3.3.2 Factors Affecting SA

Related to the error classification above, research has observed a number of factors that affect SA. Typically, these are factors that can lead to decrements in SA.

Attention. Limits of attentional capacity clearly affect perception of the situation (Endsley, 1995a). Although some pre-attentive processing can be performed in parallel, people generally must direct attention serially to perceive and process information. Thus, the

attentional limit has as much to do with the timing of information as the amount of information. Because naval C2 is a fast paced environment, with multiple channels of data, decision makers have difficulty attending to all incoming data. Information sampling strategies (Wickens, 1992a, cited in Endsley, 1995a) can help operators deal with multiple, rapid streams of information.

Working Memory. WM is crucial for mental computation but is severely limited (Baddeley, 1990). Thus, WM is a major limiting factor on both perception of data and the incorporation of data in a mental model. Large amounts of data can exceed a person's ability to retain and integrate the data, resulting in an incomplete mental model of the situation. Scheduling tools (aids to organizing data gathering activities and tasks) can improve the use of WM by helping the person look for and process information in an organized fashion (Campion, Brander & Koritas, 1996).

Stress. A certain level of stress is necessary for optimal cognitive functioning. Too little stress results in boredom and loss of attention. Too much stress, however, can be even worse. High levels of stress are associated with a restriction of attentional focus and concomitant losses of SA. As stress increases beyond some ideal moderate level, a person's attentional field becomes increasingly limited to just a few central elements. This poses the risk of premature closure, of accepting a mental model of the situation without fully exploring all the available data.

Expertise. Experience can improve SA by enhancing attentional and WM skills. Experts are generally better able to assess situations can create more detailed and accurate mental models (Randel et al., 1996). As discussed in Section 3.1.3, experts make more efficient use of memory, which gives them an advantage both in perceiving relevant information and integrating data in their mental models.

3.3.3 Recommendations

- Design decision support to help decision makers avoid SA errors at all three levels of SA.
- Develop SA support in the areas of memory aids, attention enhancement, recognitional strategies, schemata, and simulation tools.

3.4 Expertise and Training

3.4.1 Stages of Expertise

Anderson's (1995) theory of expertise, reviewed in Section 2.4.1, had two main premises. The first is that expertise develops in a series of three stages, from cognitive, to associative, to autonomous (Anderson, 1995, Ch. 9). These stages encompass changes in the underlying representations and processes used to guide performance. Initially, performance is guided by declarative knowledge of the domain and explicit rules. After some practice, the person acquires associative rules that allow the person to perform on the basis of cue-response pairings. Finally, these associative rules become incorporated in more complex production systems that can be performed automatically.

The second premise is that experts organize their memory and domain knowledge better than novices (Anderson, 1995; Endsley, 1997; Flin, 1998; Lipshitz & Shaul, 1997). In particular,

experts have created more extensive schemata that are based on the conceptual structure of the domain problems. This knowledge base allows experts to identify problems more accurately and retrieve more relevant solutions than novices.

The first premise appears not to have been tested in the naval C2 domain. A great deal of general support, however, has been obtained in other domains, such as formal problem solving, chess, physics, reading, and other skill domains (see Anderson, 1995). One piece of evidence is the gradual, but often accelerating, effect of practice (Anderson, Fincham, & Douglass, 1997; Anderson & Fincham, 1994). Domain-relevant practice is necessary to make the transition from a declarative understanding of a domain to the associative knowledge necessary for more rapid and accurate performance. Still further practice is needed to automatize performance (Anderson & Fincham, 1994).

A second piece of evidence is the increasing proceduralization of knowledge with growing expertise. Expertise is not simply a matter of acquiring more knowledge or even organizing knowledge more effectively. It is a matter of changing the form in which knowledge is represented. Anderson (1982, cited in Anderson, 1995), for example, examined expertise in solving geometry problems. He observed that initially novice problem solvers generated elaborate verbal descriptions of the problems as they searched for rules to prove geometric theorems. Over time, problem solvers needed less description of the problem. Expert problem solvers were able to quickly categorize the type of problem and identify the appropriate theorems. Similar findings have been found in expertise in physics (Sweller, Mawer, & Ward, 1983, cited in Anderson, 1995), where experts are able to immediately write formulae in terms of the constants and values of the problem at hand. In contrast, novices must engage in an extensive conceptual analysis of the problem to identify appropriate formulae and then substitute values to address the problem.

The process of proceduralization is central to intuitive theories of decision making. Experts can employ feature matching and other recognitional strategies because they have organized their memories in a way that facilitates rapid access to problem templates that describe common types of problems and the kinds of solutions that can be applied.

Although proceduralization marks the most advanced stage of expertise, there are also changes in the organization of knowledge that facilitate performance. Some of these changes have to do with the acquisition of knowledge through practice. Experts, for example, have stored a great deal of practical knowledge about their domain (e.g., de Groot, 1965). Other changes pertain to the bases for organizing knowledge. Experts organize their knowledge around the important concepts and principles of the domain (e.g., Lesgold et al., 1988). Thus, the schemata or templates used by experts in problem solving contain knowledge of the meaningful relations between problem elements. Schemata of novices tend to contain small, disconnected units of information that must be consciously linked when solving a problem (Glaser & Chi, 1988).

Overall, better, more meaningful organization of knowledge allows experts to spend less time representing problems and searching for workable solutions, and more time implementing solutions. Table 3.3 (adapted from Sternberg, 1996, p. 374) summarizes empirical findings regarding memory and cognitive processes of experts and novices.

EXPERTS	NOVICES
Have large, rich schemas containing a great deal of declarative knowledge about domain	Have relatively impoverished schemas containing relatively less declarative knowledge about domain
Have well-organized, highly interconnected units of knowledge in schemas	Have poorly organized, loosely interconnected, scattered units of knowledge
Spend proportionately more time determining how to represent a problem than in searching for and executing a problem strategy	Spend proportionately more time searching for and executing a problem strategy than in determining how to represent a problem
Develop sophisticated representation of problems, based on structural similarities among problems	Develop relatively poor and naïve representation of problems, based on superficial similarities among problems
Work forward from given information to implement strategies for finding unknown	Work backward from focusing on unknown to finding problem strategies that make use of given information
Generally choose a strategy based on elaborate schema of problem strategies; use means-ends analysis only as a backup strategy for handling unusual, atypical problems	Frequently use means-ends analysis as a strategy for handling most problems; sometimes choose a strategy based on knowledge of problem strategies
Schemas contain a great deal of procedural knowledge about problem strategies relevant to domain	Schemas contain relatively little procedural knowledge about problem strategies relevant to domain
Have automatized many sequences of steps within problem strategies	Show little or no automatization of any sequences of steps within problem strategies
Show highly efficient problem solving; when time constraints are imposed, solve problems more quickly than novices	Show relatively inefficient problem solving; solve problems less quickly than experts
Accurately predict the difficulty of solving particular problems	Do not accurately predict the difficulty of solving particular problems
Carefully monitor own problem-solving strategies and processes	Show poor monitoring of own problem-solving strategies and processes
Show high accuracy in reaching appropriate solutions	Show much less accuracy than experts in reaching appropriate solutions
When confronting highly unusual problems with atypical structural features, task relatively more time than novices both to represent the problem and to retrieve appropriate problem strategies	When confronting highly unusual problems with atypical structural features, novices take relatively less time than experts both to represent the problem and to retrieve problem strategies
When provided with new information that contradicts initial problem representation, show flexibility in adapting to a more appropriate strategy	Show less ability to adapt to new information that contradicts initial problem representation and strategy

Table 3.3 – Differences between Experts and Novices (from Sternberg, 1996)

Some findings regarding expert knowledge have been replicated in the naval C2 domain. Lipshitz and Shaul (1997), for example, performed a study investigating decision making by

Israeli Defense Force patrol boat commanders. Participants performed a scenario in which they commanded a team of patrol boats and attempted to identify and intercept enemy vessels. Some participants were active patrol boat commanders and others were novice non-commissioned officers with no command experience. Lipshitz and Shaul (1997) found five major differences between their expert and novice participants, consistent with the literature on expert-novice differences:

- Experts collected more information about the situation before making a decision.
- Experts engaged in a more efficient search.
- Experts identified situations more accurately.
- Experts identified the correct COA more often than novices.
- Experts communicated more frequently and elaborately with friendly units.

Thus, experts systematically obtained the information necessary to characterize the situation, then immediately retrieved a workable COA. Overall, Lipshitz and Shaul's (1997) results indicate that their expert participants had better mental models of the scenarios that allowed then to identify appropriate COAs.

3.4.2 Training Implications of Expertise

The main implication of research on expertise is that training should provide large amounts of practice. Expertise involves acquiring large numbers of specific problem templates and solutions (Anderson, 1990). Because the expert ultimately stores this knowledge in a procedural form, it is essential that people directly experience domain problems by doing actual tasks. In this way, they can learn about all aspects of the domain while learning how to implement solutions.

Of particular importance in training is linking actions and solutions to schemata (Lipshitz & Shaul, 1997). During the second stage of expertise, a person forms associative memories that allow him or her to identify the appropriate actions given the situation. Explicit training can facilitate this process by identifying the critical cues and factors for categorizing situations and matching them to a template.

A second implication is that practice should be concrete and realistic but highlight the conceptual structure and principles of the domain. The conceptual structure consists of the critical cues and associations that link actions to templates and forms the basis of the way an expert identifies situations. Thus, training should help people look past the perceptual surface structure of the problem (i.e., the raw data) and identify meaningful aspects and relationships in the situation.

A third implication is that trainees need different forms of training at different stages of expertise. Initially, the trainee will function at the cognitive level. Training must, to some extent, focus on providing declarative knowledge of the domain in an analytic format that the trainee can understand and incorporate in his or her growing understanding of the domain. The training, however, must also help the trainee move to the next stage by providing practice in categorizing situations and associating responses. At the second stage, declarative information becomes less valuable and the focus of training should be on expanding the associative knowledge base of trainees. This can be done through extensive practice and explicit training in critical cues. Finally, as the trainee moves into the procedural stage, the trainee will require less formal instruction other than to find and remediate errors. What will

be crucial is realistic simulation and practice to help the trainee apply what has been learned in real settings.

Anderson (1990) advocates a componential approach to training, in which material is broken into related sets of elements. These elements are taught individually through mastery learning techniques. With mastery learning, trainees study one component at a time, while their learning is monitored. Trainees do not advance until they have mastered the component.

3.4.3 Potential Problems

At least two factors can limit the expression of expertise. These constitute potential problems that should be dealt with either in training or through the use of a DSS.

3.4.3.1 Transfer of Skill

Typically, experts exhibit narrow applicability of their skills (Anderson, 1990), functioning as experts only in a constrained problem domain. When asked to perform in a similar or related domain, experts may perform no better than complete novices. This problem arises not because expertise lacks applicability to other domains but because experts fail to see how to apply expertise. Expertise itself does not help a person see relationships between related domains and apply expert skills in new situations.

One way to deal with this problem is through use of DSSs that identify principles and features of one domain that are common to another. Simply pointing out analogies between problems in one domain and problems in another can dramatically improve the use of existing knowledge and expertise (Gick & Holyoak, 1983).

Another way to deal with this problem is to deal with it explicitly in training. As discussed below in more detail, cross training of C2 team members on the roles and responsibilities of all other members can enhance the functioning of a team. It can also allow an expert in one specific aspect of C2 to solve problems in related areas.

3.4.3.2 Mental sets

Related to the problem of applicability is the problem of mental set. Mental set refers to the rigid application of expert knowledge at the expense of creativity and flexibility. Experts can sometimes become fixated on applying the kinds of solutions they have learned in their domain of expertise to all situations, even when those solutions are not appropriate (Wiley, 1998). This can lead to cases where novices actually outperform experts because the novices have less difficulty in generating creative solutions to novel problems.

Training should always include some component that deals with recognizing when expert knowledge can be directly applied (Wiley, 1998). This corresponds to the common step in tactical problem solving of identifying whether there is a problem or not (e.g., Gilhooley, 1989; Miller at al., 1992). This step should include the additional judgment of whether the problem falls under the scope of previously learned or experienced problems or whether it is a novel problem requiring a new kind of solution. In addition to training, DSSs should offer indications of the overall type or structure of the situation to help the decision maker decide how his or her experience applies (Wiley, 1998).

3.4.4 Approaches to Training

This section reviews training programs and suggested training practices for naval C2.

3.4.4.1 Critical Thinking Training (CTT)

CTT was developed on the basis of intuitive theories of decision making. The goal of CTT is to provide training in critical thinking skills to enhance decision making (Johnson & Cannon-Bowers, 1996). It also focuses on providing realistic experience and practice to help trainees learn how to apply decision skills in the field (i.e. learning-by-doing; Section 2.4.3.1).

Freeman and Cohen (1996) developed the STEP (Story, Testing, Evaluation, and Plan) procedure based on the Recognition/Metacognition model (see Cohen et al., 1997). The goal of STEP is to help trainees build situation models or templates that will form the basis of expertise. Training is highly context dependent and practically oriented. Trainees learn by solving scenarios and analyzing the nature of scenarios. STEP consists of four instructional units:

Story. Build a story to explain the scenario situation, with past and future events that will be true if the story is correct.

Testing. Test the story for conflicts by comparing story-based expectations to what is known or observed.

Evaluation. Evaluate one's assumptions; if there is too much disconfirming evidence, begin the cycle again.

Plans. Develop plans to deal with the contingency that one's current best story is wrong.

Cohen et al. (1997) identify a number of benefits of this approach, notably that it facilitates development of a "big picture." Overall, the STEP training regimen teaches trainees how to solve tactical problems by doing what current experts do. To do so, it makes use of real scenarios gathered from active personnel and implemented in the DEFFT computer-based simulation, giving STEP a high degree of operational relevance.

Freeman and Cohen (1996) empirically validated the STEP approach. They compared decision making in realistic scenarios of personnel trained by the STEP approach to that of personnel trained by conventional methods. In particular, they examined the impact of training on decision processes, such as generating alternative explanations while assessing the situation, and decision outcomes (i.e., the quality of decisions). With respect to decision processes, Freeman and Cohen (1996) found that personnel trained by STEP:

- Identified more conflicting evidence in scenarios.
- Generated more explanations.
- Gave more arguments supporting assessments.
- Produced more alternative assessments.

With respect to decision making performance, they found that STEP training:

- Increased the agreement between trainee decisions and those of SMEs.
- Increased consensus among trained officers.
- Improved accuracy of situation assessment.

Overall, STEP seems to be a good technique to provide the kind of realistic practice needed to develop expertise.

Another intuitive-based technique, developed by Pliske, McCloskey, and Klein (1998), focuses, like STEP, on explicitly identifying the conceptual structure of problems. Their approach makes use of tactical simulations that present the trainee with a problem requiring assessment and decision. The goal again is to help trainees learn the kinds of situations they will face in the field and the kinds of solutions that can be applied. To help trainees clearly identify the critical cues and understand good solutions, trainees engage in a number of *critiquing exercises*.

Decision making critique. In this exercise, trainees respond to a set of questions designed to identify critical decisions made during the exercise. They explore cues that they could have identified earlier, assess their mistakes, and discuss their uncertainties. This exercise helps the trainee identify what went well and what went wrong in the scenario.

Decision Requirements (DRs) exercise. This exercise identifies DRs in the scenario. It helps trainees gain familiarity with the kinds of decisions they will face in real situations. Trainees determine why these decisions were the most challenging in order to gain an understanding of the problems they will face.

PreMortem exercise. During this exercise, trainees identify key vulnerabilities in a plan. One trainee develops a plan and the training group is told that the plan failed. The group spends a few minutes writing down reasons why the plan failed then discuss reasons put forward by each group member. This exercise helps trainees uncover flaws in plans and learn self-criticism.

Commander's intent exercise. This exercise improves trainees' ability to communicate intent. Trainees describe a solution to the scenario in the form of a set of orders to subordinates that convey the trainee's intent. A training facilitator identifies plausible events that could interfere with the plan. The trainee writes down how he or she expects the subordinates to react, while the subordinates write down how they would react. These responses are compared in a group discussion to note differences and identify ways to make intent clearer.

3.4.4.2 Formal Versus CTT Approaches

CTT methods contrast with more traditional, formal training methods. Formal methods (Fallesen & Pound, 1998):

- Are highly structured.
- Facilitate explicit learning.
- Focus on prescribed rules and procedures.
- Derive teaching points from analytic theories and models.

CTT methods, on the other hand (Cohen et al., 1997; Fallesen & Pound, 1998):

- Have less formal structure.
- Facilitate implicit learning.
- Focus on practical strategies of experts.
- Focus on recognizing situations.
- Derive teaching points from intuitive theories.

Fallesen and Pound (1998) developed a hybrid approach that combines elements of formal and CTT techniques. Like CTT, this approach makes use of scenario-based instruction so that trainees learn in the context of realistic problems. It also focuses on a three-stage decision making strategy:

- Form SA.
- Identify goals and plans.
- Translate plans into actions.

Trainees engage in relevance checking at each stage. The trainees ask questions of their own reasoning processes, seeking to determine whether they have reached the correct conclusion. Formal aspects of the training include lecture-based declarative descriptions of the meaning, relevancy, and rationale of their decisions from instructors, which help trainees develop a conceptual understanding of the domain trainees receive. The approach also makes use of formal concepts in giving feedback.

A hybrid approach could be particularly valuable. Although tactical decision makers most often used intuitive strategies, they do occasionally use analytic strategies (Pound & Fallesen, 1994, cited in Fallesen & Pound, 1998). Formal training approaches may be better suited for training analytic strategies than CTT techniques. Also, a hybrid approach fits with the stages of expertise discussed above. Trainees require some structured, formal instruction, at least initially, to build their declarative knowledge base.

3.4.4.3 Multimedia

As discussed in Section 2.4.3.4, multimedia and learning-by-doing techniques can be very powerful in training. Learning-by-doing, in particular, may be particularly well suited for tactical training because it allows trainees to experience the consequences of different COAs and develop an understanding of good and bad solutions to problems (Anzai, 1984).

Radtke and Frey identified a number of ways that multimedia can be used to train Tactical Decision Making (TDM). In particular, multimedia can:

- Provide trainees with a representation of situation cues.
- Highlight key cue patterns to aid trainees to isolate important information from noise and distractions.
- Present tactical knowledge in logical and accessible formats, providing organization for trainees' mental models.
- Present tactical knowledge in a form that trainees can recall when needed (i.e., vivid imagery, mnemonics).
- Demonstrate skills and procedures needed to select, plan, and implement actions.
- Facilitate practice of skills and knowledge.
- Provide feedback of effectiveness compared to standards of performance.
- Illustrate the relationship between events in the external environment and cues observed in the command centre.

Overall, multimedia seems crucial for creating the realism necessary for experiential learning. Table 3.4 contains a more detailed list of the benefits multimedia training should provide (Radtke & Frey, 1996).

Criteria of Instructional Effectiveness
• The information in the presentation is well organized and is presented in a logical order
• The presentation provides personalized feedback on the learner's performance
• The learner can review material as often as wanted
• The learner can navigate in the presence without becoming lost or stuck
• The learner sets the pace at which material is presented
• The presentation uses visual illustrations to make the ideas and facts being taught clearer and more concrete
• The material is presented in concise and manageable "chunks"
• The learner can explore a topic of interest in more detail
• The presentation forces the learner to think about the materials being presented
• The presentation holds the learner's attention by making the material interesting and entertaining
• The presentation provides opportunities for the learner to step back and think about how the small details all fit together
• The presentation uses vivid visual images to help the learner remember important points or ideas
• The presentation uses animation and graphics to show ideas and processes that otherwise could never be seen in the real world
• Key points are presented in several different ways
• Important points are clearly identified with visual or sound cues

Table 3.4 – Criteria for Effective Multimedia Instruction (from Radtke & Frey, 1996)

3.4.4.4 Team Training

Most training approaches consider the individual but tactical decision makers will work in teams. Thus, it is important to consider team training. Certainly, some part of training should address teamwork issues, such as coordination, communication, and team roles (Cannon-Bowers et al., 1995; Radtke & Frey, 1996).

One aspect of teamwork that researchers have addressed is the need for team SA or shared mental models (Rouse et al., 1992. One way to facilitate shared mental models and the ability to form expectations about teammates is through *cross-training*. Cross-training consists of team members learning to perform each others' duties and acquiring knowledge of each team member's role and responsibilities (Rouse et al., 1992; Volpe et al., 1996). This can be done by having each member serve at each position in a series of training scenarios.

Cross-training should result in greater common knowledge and experience within a team. It should also facilitate communication and the efficient division of tasks (Rouse et al., 1992). Thus, Rouse et al. (1992) argue that team members who have developed shared mental models of the teamwork and taskwork will:

- Be more accurate in predicting behaviours of teammates.
- Generate similar explanations of situations.

- Require less overt planning time.
- Will communicate less but maintain performance.
- Will request information less frequently.
- Be more resilient to stress effects and need less explicit coordination.

Empirical evidence supports these hypotheses. Volpe et al. (1996) compared a cross-training regimen with a conventional, individual training regimen. Student B-29 aircrew teams were assigned to either cross-training or standard individual-oriented training. The team tasks consisted of flight simulations and aircrew planning. Volpe et al. (1996) measured teamwork processes and individual performance as well as communications frequency. They found that:

- Teams with cross-training performed better on measures of teamwork.
- Teams with cross-training were more efficient in communication, requesting less information, volunteering more information, and reducing irrelevant communications.

Thus, cross-training seems to help teams better anticipate one another's needs and coordinate activities, including communication.

3.4.5 DSSs and Training

Placing DSSs aboard the HALIFAX class raises additional training issues. C2 teams will need training in how to use the DSS but also training in how to make decisions when using the DSS. Training should allow teams to experience the specific systems that they will use in the field. Moving some training to the shipboard environment is one way to accomplish this and potentially improve training effectiveness and reduce training costs (Johnson & Cannon-Bowers, 1996). This increases the fidelity of training to real conditions and allows training to be rapidly adapted to changing conditions.

3.4.6 Recommendations

- Provide extensive practice in realistic operational settings so that training covers all anticipated tasks and situations, helps trainees identify situations in which expertise can be applied, and increasingly emphasizes performance over explanation as the trainee gains experience.
- Design training to support trainees' levels of experience by initially focusing on declarative knowledge and the acquisition of a knowledge base then shifting to focus on proceduralizing knowledge and detecting and remediating mistakes
- Adopt componential training that allows trainees to move on only after mastering the previous component.

3.5 Teams

Although a fair amount of research has been done on teamwork and team-related variables (cohesion, communication, etc.), the scope of the literature review precluded a thorough review of empirical results related to teams. Teamwork may be a topic worthy of a separate literature review. Nevertheless, the theories reviewed in Section 2.5 indicate the kinds of factors and processes that influence the effectiveness of teams.

3.6 Human-Computer Interaction

Most work on HCI issues has been practical in orientation, thus limiting general empirical results. This section, however, will review several lines of study that contribute to general design issues.

3.6.1 C2 and HCI

C2 tasks place a large number of information processing demands on personnel. Howell et al. (1993) analyzed C2 tasks from the perspective of HCI to identify specific cognitive functions that must be supported by an interface. They drew on four sources to classify task elements:

- Previous taxonomies of supervisory control functions.
- Documentation of C3I systems.
- Direct observation of exemplar systems.
- Current theories of human cognition.

Their analysis revealed a set of 18 major operations, which are listed in Table 3.5. This set is not an exhaustive analysis of an OR aboard a particular vessel and does not address the many highly specific functions for each position in the OR. Nevertheless, these operations serve as the basis for generating HCI principles to be used in the general design of interfaces (Howell et al.'s, 1993, guidelines will be reviewed in Section 7).

In addition to serving as a basis for processes, these categories can be used by designers to frame prototyping and user evaluation in rapid prototyping methods. They can also be used in Cognitive Task Analysis (CTA) to identify HCI issues prior to developing design concepts.

The list of operations also provides a basis for determining how cognitive processes can be affected by HCI design and vice versa. There is no hard-and-fast methodology to do this but HCI and cognitive theories provide some basis. Howell et al. (1988) summarize a number of observations regarding the effects of display and HCI characteristics of C2 systems on cognitive processes but two points in particular bear discussion.

First, the HCI must accommodate limits of human attention, WM, and information processing. The presence of irrelevant information interferes with the extraction and interpretation of information (Howell et al., 1988). Furthermore, information that is logically related to the task can nevertheless be irrelevant if it is presented in such a way that the user does not integrate or process it in a task-relevant manner. In this case, the information simply becomes more distracting data. Thus, two important principles of HCI are:

- Identify and present only relevant information (i.e., data).
- Partition information in a way that indicates logical relations between data.

Both these principles rely on task analysis and CTA to create representations of the functions and processes of the task and the cognitive strategies of the people doing the task. Together, these representations create the logical divisions between task components and information requirements.

Major Cognitive Operations Performed on Information in C3I Tasks
Simple Extraction
Read-out
Identify/recognize
Locate
Complex extraction
Discriminate/compare
Filter/ignore
Perceive/interpret pattern
Correlate
Monitor
Process
Remember
Estimate
Calculate
Integrate/organize/aggregate
Evaluate
Generate/create
Choose/decide
Manipulate
Command system
Multiple Operations
Complex interaction

Table 3.5 – Major Cognitive Operations Needing Interface Support (from Howell et al., 1993)

A second finding is that displays can induce or even require either serial or parallel processing. A display that presents pieces of data one bit at a time, for example, favors serial processing, whereas a display that presents all data for a limited amount of time favors parallel processing. The choice of which kind of processing to facilitate is crucial because some cognitive processes are serial and others parallel. Perceptual and some attentional and motor processes are parallel and can proceed with large amounts of data (e.g., Sternberg, 1996, pp. 211-216). To take advantage of the processing capacity and speed of these processes, the display should not only allow but actively facilitate parallel processing. Higher level cognitive functions, however, are generally serial and much slower than perceptual processes. These processes require a more serial interface, one that either spaces data over time or maintains data indefinitely for the user. Because most decision making processes are serial, DSSs will need a highly serial interface.

3.6.2 Display Issues

How information is presented can change the user's decision making strategy (Jarvenpaa, 1989; Johnson, Payne & Bettman, 1988, cited in Adelman et al., 1993). The display format certainly affects memory and attention demands but, as mentioned above, display characteristics can also facilitate or impair certain kinds of cognitive strategies. This section discusses two HCI aspects particularly relevant to this issue.

3.6.2.1 Graphic Versus Text Displays

Problem solving and decision making depend on the representation of the situation. The principle of *cognitive fit* demands that, for optimal performance, the user engage in the same or similar cognitive processes to represent information in the interface as engaged in understanding and solving the problem (Vessey, 1991). This means that the person need not transform the initial representation of the interface to extract the information into a form appropriate to the problem. Cognitive fit increases speed and accuracy of problem solving. It reduces the amount of cognitive processing required by a task because the display presents only the data pertinent to the decision of the moment.

Two basic ways of depicting data are through text or tables of numbers and through graphics, including diagrams and graphs. Text and tables are verbal, discrete, and analytic representations (Vessey, 1991). They present information in symbolic form so that precise values of variables are well-defined but relationships between variables are implicit; i.e., the user must mentally compute the relation between variables. Graphics are imaginal and analog representations. They present information in a spatial form so that values of variables are less well defined but relationships are explicit (Vessey, 1991). This is especially useful for depicting changes in variables (e.g., Hutchins, 1996; Rummel, 1995). Graphics facilitate viewing large amounts of data at a glance and identifying higher-order relationships of variables but they make it difficult to identify particular values.

The implication of this analysis is that textual displays of information will facilitate tasks that require identification of precise bits of data whereas graphical displays will facilitate comparison of variables and identification of relationships. In fact, studies have observed just that. Jarvenpaa and Dickson (1988), for example, found that graphs helped people summarize data, identify trends and relationships, compare data points, and detect deviations or differences within a data set. Jarvenpaa (1989) found that even the particular form of a graphic display can affect decision making time and accuracy. Similarly textual formats that are consistent with comparison of alternatives on specific attributes can facilitate decision making (Bettman & Sins, 1979, cited in Jarvenpaa, 1989).

In a study of problem solving, Vessey (1991) contrasted performance of spatial and symbolic tasks. Spatial tasks involved manipulation of spatial relations of distance and direction. These tasks were highly relational and required comparison of spatial values. Symbolic tasks were more analytical and required manipulation of discrete bits of data. Participants performed each kind of task using either a graphic or tabular display. Graphs led to faster, more accurate performance of the spatial tasks and tables to faster, more accurate performance of the symbolic tasks.

These results have direct implications for naval C2, where many tasks entail processing relations and trends in data. Design of displays should take into account the tasks that will be supported, the kinds of information that the user will extract, and the kinds of decision strategies that will be used.

3.6.2.2 Speech Versus Keyboard Input

Traditionally, C2 systems aboard Navy vessels have used keyboard and trackball input devices. Many researchers (e.g., Lee, 1989, cited in Damper & Wood, 1995) have touted the advantages of speech as an input tool, arguing that a speech-based interface is:

- Fast.
- Natural.
- Hands-free.
- Eyes-free.
- Location-free.

Thus, a natural language interface could free the user to perform tasks more quickly and with fewer constraints (e.g., freedom to move).

Neglecting technological issues, however, there may be human factors concerns about such an input device. Damper and Wood (1995) compared the speed and accuracy of entering instructions to an electronic mail editor by keyboard versus speech interface. They varied the length and naturalness of commands to the system, testing the hypothesis that speech input would lead to better performance only when commands were long and similar to natural language. This result was obtained. Further, they found that when commands were short, keying commands was slightly faster than speaking commands and much less error prone. Damper and Wood (1995) concluded that the value of speech interfaces must be carefully evaluated on the basis of the task and the nature of commands to the system, as well as the technical feasibility of reliable speech recognition systems.

3.6.3 Embedded User Models

Overall, research in HCI indicates the need for the interface to match the information processing limits and styles of the user. Researchers have suggested using *embedded user models* in systems to enhance the degree to which the interface is adapted to the user (see Zachary & Ross, 1991).

An embedded user model consists of a behavioural and cognitive model of the typical user. It describes the goals, actions, plans, needs, and limitations of the user, taking into account also the task structure. The model serves to predict errors, information needs, and processing strategies.

Zachary and Ross (1991) employed the Cognition as a Network of Tasks (COGNET) framework to develop an embedded user model of the naval ASW. This system contains three notational tools:

- Operator's global problem representation (blackboard).
- Human-machine interactions associated with each cognitive task.
- Production rules describing information processing steps.

The model was implemented using BATON software that provides mechanisms for:

- Defining the blackboard structure.
- Entering, changing, and removing information from the blackboard.
- Defining and implementing perceptual monitors and linking them to display events in the HCI.
- Implementing individual COGNET task models.
- Implementing active HCI routines based on context triggers tied to the model.

Overall, then, the embedded user model performs two tasks. First, it actively models the cognitive processes and knowledge of the user. Second, it monitors the HCI and controls specific HCI features and functions. The embedded user model monitors the actions of the user and the information available to the user. It attempts to infer the user's current mental model of the tactical situation and cognitive processing to detect inconsistencies between the user's cognitive needs and the way information is currently being presented by the interface. The system can then adjust the interface to be in line with what the user is predicted to be doing. An intelligent interaction subsystem performs context-aiding and user support functions. For example, the subsystem can:

- Alert the user of the opportunity or need to perform or resume a task.
- Offer to perform a task automatically.
- Automate some of the lowest level manual interactions to facilitate human performance.

The success of an embedded user model depends on the completeness and accuracy with which the model can be specified. The chief problems facing this approach appear to be the difficulty of predicting beforehand all the situations and tasks that will confront the user and predicting the range of individual differences that will be encountered in the user population. Most notably, the model must accommodate differences between experts and novices because users will have different levels of experience.

3.6.4 Recommendations

- Design for cognitive fit and match display format (graphic versus text) to the task/function.
- Study the applicability and usefulness new technologies, such as speech input and embedded user models, for CPF OR workstations.

4. Summary of Theory and Empirical Results in the Operational Context

This section reviews theoretical and empirical issues from the perspective of operations. In particular, this section raises relevant issues and questions that could have an impact on the upgrade of the HALIFAX class. The goal of this section is to put the reviewed literature in a practical light. The reason for summarizing two sections together, theory and empirical results, is that two perspectives on the literature allow new insights to be revealed. In addition, this section should present the literature in a way that is of greater relevance to operational naval decision makers.

The discussion will follow three broad stages of an operation, mission planning (including planning and rehearsal), coming on watch (including watch transfer), and on watch surveillance, threat assessment, and threat response. In addition, we will initially consider the general issue of the context of operations. The context affects all aspects of decision making so that it is impossible to describe or analyze decision making in detail without knowing the context in which decision makers operate.

4.1 Operational Context

4.1.1 Single Ship Versus Task Group Settings

The objective of this literature review was to explore decision making in the context of single ship operations. One thing that has become clear, however, is that this is an unrealistic or severely limited setting in which to address naval decision making and C2. Vessels almost always operate in task groups. Frequently, these task groups are frequently multinational, joint forces operating in foreign waters and, hence, requiring international support. This consideration is especially important for the Canadian Navy, which will likely frequently participate in multinational task groups. Operating in a task group complicates the coordination of decision making activities, increases the complexity of situations, and potentially adds barriers to effective decision making. At a surface level, this means differences in language, vocabulary, procedures, and so on. At a more fundamental level, there are likely to be differences in intent, shared mental models, and goals. An important issue to resolve is how to design systems to accommodate such differences.

4.1.2 Mental Models

As discussed in Section 2.3, SA is more than just the perception of information; it is the understanding of events in relation to one's goals. In other words, SA entails creating and maintaining a *mental model* of the situation, one that allows mental simulation and prediction of future events.

This raises several questions, not the least of which is what is the mental model formed by the naval CT? To understand SA, decision making, and other important activities, we must know how each commander at each level of decision making represents the situation. But this representation is formed in a rich context, of the CT, the ship, the task group, and still higher levels of organization. Thus, mental models will differ by a number of critical factors:

- Level of naval command.
- Type of task group (uni-national or multinational).
- Presence of non-military organizations.
- Strategic aims.

This is by no means an exhaustive list but indicates the difficulty of describing the mental models that underlie SA and decision making.

A major issue, then, is how members of the naval CT develop their mental models. This issue must be addressed before one can make any progress in the design of decision support. It must be addressed, however, in relation to the various shifting contextual factors that affect how each team member understands the situation.

In addition, we must remember that decision making is a team process, relying on the coordinated action of the CT within a vessel and CTs across vessels. An important question is how command personnel can communicate their understanding of the situation to achieve common intent and understanding (see Pigeau & McCann, 1998). To date, this is an underdeveloped area of inquiry.

4.1.3 Situation Awareness

The previous discussion points to a need for support of SA. This is not a surprising or new finding but merely reinforces previous recommendations (Section 2.3.5). The discussion of Endsley's (1995; 1997) theory of SA and the nature of decision making strategies, however, does indicate a shift in thinking about SA support. It is clear that increasing access to information is not sufficient to improve SA - or even desirable.

Perception of data comprises only one level of SA. The other levels pertain to comprehending data in relation to goals and projecting events to predict future outcomes. Thus, to achieve SA, one must not only perceive and register data but process it to a sufficient degree that the data informs one's mental model and contributes to decision making processes. What is needed are tools that aid in integrating data with strategic and tactical goals and constraints (i.e., seeing data in an operationally relevant way).

4.1.4 Decision Making

The most appropriate decision making strategy will also depend on the operational context. As we have seen, analytic decision making strategies seem inappropriate to situations involving severe time pressure or impoverished data. These strategies do, however, have advantages, namely they will produce optimal solutions and can be applied even in unfamiliar situations. Intuitive decision making strategies, in contrast, are well suited to time-stressed situations because they operate quickly. They are also robust and can deal with missing or errorful data. The chief drawbacks of intuitive strategies are that they are applicable only to familiar situations and are susceptible to error if an important element of the problem is misperceived.

As a result, there is no one decision making strategy that can be selected as best for all situations. Furthermore, neither strategy seems completely suited to all phases of an operation. As we will discuss in more detail below, certain phases of an operation, such as planning and preparation, afford more time and more complete data, than do other phases,

such as implementation (activities related to surveillance, threat detection, and threat response). The practical implication is that naval decision makers will need more than one kind of decision support. At times, it will be important to support both analytic thinking and intuitive decision making at various times.

Another issue to consider in relation to support for intuitive decision making is levels at which problems can be matched to experience. Intuitive strategies rely on recognition and matching by similarity but similarity must be judged relatively. Thus, psychologists distinguish the *conceptual* and *surface* levels of problems (e.g., Gick & Holyoak, 1983). The conceptual level corresponds to the formal structure of the problem. It describes the relations between problem elements and the kinds of problem solving operations possible. Thus, responding to an air threat has a certain formal structure that governs such things as how long one has to respond, what kinds of actions one can take, and so on. What is important is that other problems, such as responding to a surface threat may share many of the same formal elements, making the two problems similar in some respects. The surface structure corresponds to the perceivable characteristics of the problem elements, the specific kinds of aircraft, ships, terrain, and so on of the specific problem under consideration. These elements are distinct from the formal structure of the problem but are the most immediate and easy to perceive features of the problem.

One issue that should be addressed is how expert decision makers use conceptual and surface information. The distinction between conceptual and surface similarity opens possibilities for different classes of intuitive decision making errors. In particular, decision makers who rely on surface features of the situation could fail to accurately recognize the situation. The danger is that the decision maker would relate the situation to some past experience that had surface similarity but was unrelated at a conceptual level. In this case, the decision maker might believe that he or she had identified a workable solution when in reality it was not appropriate. It is also possible, however, that decision makers who neglect the surface level might have difficulty. Surface level features can help one recognize past experiences that have both surface and conceptual similarity to the current situation. If surface features are neglected or if the situation is superficially dissimilar, a decision maker might fail to match the current problem to a relevant past solution.

4.2 Mission Preparation

This stage of operations includes the planning and rehearsal that precedes any action. It may be somewhat misleading to refer to this as a stage because planning and rehearsal activities are not confined to some discrete portion of time. Instead, they are ongoing activities carried out continuously during operations. Nevertheless, it is useful to distinguish mission preparation from activities of the watch and consider how the one affects the other. Activities on watch, for example, are to a great degree guided by expectations, mental models, and/or procedures established in advance during mission preparation.

4.2.1 Decision Making Strategies

As suggested above, there may exist a discrepancy between the decision strategy suited to planning and the strategy suited to action. Because planning typically takes place under (relatively) light time pressure and sometimes (if not always) with relatively large amounts of

high quality data, analytic strategies seem better suited to this stage. When planning, decision makers are more likely to have the time and access to resources (possibly shore-based) needed to clearly define the problem space and consider multiple COAs. Also, there is greater emphasis on determining the best possible solutions rather than merely satisfactory ones. In contrast, the severe time pressure and uncertainty of action states (e.g., responding to a threat) make analytic strategies less tenable. Instead, intuitive pattern matching strategies seem best suited in these cases.

This discrepancy in and of itself is not a bad thing. In fact, there is the possibility for synergy between analytic and intuitive strategies that could enhance overall decision making success. Specifically, analytic strategies could be used in planning to generate high quality COAs to anticipated problems so that these COAs can be rapidly selected and implemented by intuitive, recognition-based decision making processes during threat response. In this way, the benefits of analytic decision making can be indirectly imported to intuitive decision making.

One problem with this, however, is that planned responses are not helpful unless the decision maker recognizes the relation of the current situation to the planned contingency. This problem must be dealt with every time a decision maker makes a decision because the match between the current situation and experience will never be exact. Although members of the CT will likely have no difficulty recognizing broad classes of pre-planned situations, the specifics of any tactical situation will be unique to some degree. Due to the nature of analogy and intuitive decision making (e.g., Gick & Holyoak, 1983; Klein, 1997), a decision maker may fail to see the connection between the current situation and pre-planned responses or the degree of modification required. This failure occurs when the situation is not exactly the same as the situation anticipated in planning – even if the planned contingencies are similar or even formally equivalent to the current situation.

Thus, a major area of opportunity for the design of decision support is improving the connection between planning and action. Tools that help the decision maker link on-going situations to pre-planned responses would improve the effectiveness of planning. Such decision support could take a number of forms, from memory aids to improve recognition and recall of plans to inference tools that monitor sensor data and indicate when planned contingencies are present (see SABER and TADMUS DSS, Section 3.2.7). In addition, interface design should minimize the conversion effort required to transmit the results of decision making from one individual to another.

4.2.2 Interface Design for Planning

Related to this issue of support for implementing planning, attention should be given to the user-interfaces of decision support at the planning and action phases. Far more effort has been devoted to the design of workstations for crew on watch (i.e., the OR), whereas little has been devoted to computer equipment for planning. Currently in the Canadian Navy, onboard planning appears to be done off watch using commercial software applications or pencil and paper.

The question that needs to be addressed in future research is what kind of interface will promote greater transfer between planning and action. The goal is to improve interoperability across time and within and across ships. Taking into account the likelihood that Canadian ships will operate as part of a multi-ship, multi-service, multinational force with command duties assigned between as well as within ships, this is more than a within-ship challenge.

Interface design, however, can contribute to much more than the transfer of planned responses to implementation. It can also help improve planning activities themselves. The Ecological Interface Design principle advocates an approach aimed at providing the decision maker with an understanding of the system (in this case, the tactical situation including one's own ship and resources) (Vicente & Rasmussen, 1992). In doing so, the interface gives the decision maker a better understanding of the underlying problem space and greater insight into different demands.

How interfaces are designed is of paramount importance. The concepts of rapid prototyping and iterative design recognize that designers may not be able to arrive at a workable design solution except as part of an iterative design cycle with user involvement and realistic scenarios.

Analytic decision making depends heavily on a complete and accurate representation of the problem space. The better able the interface is to help in the creation of this representation, the better it is able to assist the decision maker in identifying specific steps towards a solution. Thus, in the planning stage, the interface should be less concerned with mimicking OR workstations or providing extensive data. Rather, it should be concerned with providing tools for understanding potential situations at a conceptual level.

To do this, the design of an interface must itself be based on a conceptual model of naval tactics and C2. Given such a model, interface design for planning activities becomes an exercise in determining how best to convey that model. A great deal of research has been done to develop theories of C2 (e.g., Flin, 1998; Pigeau & McCann, 1998; Stevens et al., 1996) but there has been little agreement on a suitable conceptual model. What is needed is a commonly accepted model that defines C2 and describes how it works. Developing such a model will be a prerequisite of effective interface design.

One issue that must be taken into consideration is that even if a conceptual model is developed, other forces may not adopt it. This is particularly likely when considering multinational or joint service operations. Thus, naval CTs will inevitably find themselves in the situation of planning operations with forces operating on a different model of C2. Some means of accommodating differences will have to be developed to reduce the potential misunderstanding created by this situation. Again, this is especially important for the Canadian Navy because it will generally coordinate with larger and multinational forces of politically influential nations, necessitating that Canadian vessels "fit in" to their C2 processes.

On the positive side, this problem may be mediated by having a conceptual model that will allow one to determine potential communications problems. In essence, the model will make it clear *where* and *why* forces fail to understand one another, if not how to resolve the misunderstanding.

4.2.3 Planners and Actors

Mission planning involves interaction between shore-based planning teams and planning teams afloat (for the task group and for individual vessels). Implementation, however, is ultimately carried out by the CT afloat. One key to the success of implementing plans effectively is the communication and understanding between planning and implementing teams.

Unfortunately, there may be discrepancies between the perspectives of these two teams. For example, operational planning teams may have more actual operational experience than those at sea. That experience, however, may have been obtained using different technologies and weapon systems, in different operational settings (theatre), and so on. It may also have been years since planners ashore last had operational experience. Planning teams afloat will be closer to the action and current operational context but may lack the breadth of experience of planning teams ashore. The practical implication is that different planning teams may have difficult understanding one another due to differences in how they conceive of C2 in the operational setting.

This raises a question regarding the effectiveness of planning. As noted above, actors are likely to rely on intuitive, recognition-based strategies to retrieve suitable COAs during time-stressed events. If they have not themselves participated in generating pre-planned responses, it will be much harder for them to match the situation to planned contingencies and retrieve the pre-planned response.

Furthermore, planners who are distant from the operational context may operate on an inadequate conceptual model. In particular, constraints imposed by modern technology or military practice may be overlooked or downplayed to the detriment of mission planning.

No relevant literature was obtained for this review that even addressed this issue. However, if we accept Pigeau and McCann's (1998) argument that C2 is based on the importance of establishing of common intent (implicit or explicit), this issue becomes critical. Differences in the background and training of decision makers will serve as a barrier to understanding. Therefore, it is necessary to learn more about what barriers exist and how they affect effective C2 and tactical decision making.

4.2.4 Expertise and Training

The very aspects of expertise that make it possible to respond quickly and accurately with intuitive decision making strategies can actually be barriers to effective communication and team performance. Anderson's (1995) theory of expertise indicates that knowledge becomes increasingly proceduralized with experience. That is, people move from an initially declarative, conceptual understanding of a domain to a schematic, performance-oriented understanding. This facilitates action but removes knowledge to some extent from conscious consideration and makes it hard to verbalize what one knows. This makes expert mental models resistant to reinterpretation and translation because they are procedural.

Furthermore, expertise is highly context-dependent, being acquired through practice and exposure to numerous realistic problems. An expert is someone who is readily able to recognize a situation and retrieve an appropriate solution. The downside of this is that expertise is not always transferable to new situations. Even if one is able to step back and analyze where one's expertise came from, one may not be able to sever the contextual links that allow access to schematic knowledge. Problems that are conceptually similar to ones previously experience can appear radically different if the context is not familiar.

An issue that needs to be addressed to promote better integration and interaction of CTs is how to overcome the effects of proceduralization and promote exchange of understanding and intent. This could entail tools to aid members of the CT to interpret the mental models of

others but could also entail tools to help experts consciously interpret their *own* mental models and prompt metacognitive strategies to overcome error inducing conditions.

4.2.5 Teamwork

One limitation of the literature on teamwork is that it has tended to focus on relatively small teams operating in the same space, commonly face-to-face. This may be appropriate when considering the functioning within an individual CT but it says nothing of how that CT interacts with other CTs in the larger military structure. In particular, research is needed on the establishment of common intent across teams and how cohesion within a team affects its communication across teams.

Although a fair amount of research has examined the functioning of teams, including CTs (e.g., Cannon-Bowers et al, 1995; Fleishman & Vaccaro, 1992), little work has been done on the interaction of multiple teams. By this, we are not referring to teams within the context of a larger, hierarchical organization but teams formed and maintained in separate organizations. This situation becomes increasingly important when considering the interaction of CTs of different nations or cooperation of military and non-military organizations. It is not clear that the same principles that govern effective teamwork within one organization will generalize across organizations.

An issue that remains to be resolved is the nature of team SA. Two contrasting views have emerged in the literature. The first is that team SA corresponds to shared knowledge across all team members. In this view, individuals need to share information and achieve a unity of understanding so that each team member has essentially the same mental model of the situation. The second view is that team SA corresponds to the degree to which each team member has the knowledge necessary to perform his or her role within the team, regardless of whether that information is shared with other team members. This view focuses on the appropriate placement of information within the team and the formation of individual SA sufficient to perform individual team-related tasks.

Determining which view best accounts for team behaviour in the OR is crucial to the design of decision support for the CPF. The first view implies a need for common communication and information structures so that all team members can build the same mental model. Essentially all members must be involved in reaching an implied agreement in assessment of the situation and all members must be kept informed of changes and new information. The second view implies that each team member has unique information needs related to his or her role. Hence, there must be communication and information structures designed to deliver that information appropriately and without interference. This view suggests that common communication structures could actually hurt team SA by burdening members with information not relevant to their tasks.

Nothing found in the existing literature resolves this issue. Thus, a research program to investigate team SA would prove valuable to the upgrade of the CPF.

4.3 Coming on Watch

4.3.1 Situation Awareness

Coming on watch poses serious problems for SA that, so far, have not been addressed in the literature. SA arises from a continuous monitoring and integration of data to form a meaningful mental model of the situation (Endsley, 1995). Likewise, a mental model is used to determine what information to seek in the situation. When members of a CT come on watch, they may have been "out of the loop" for several hours, making their mental model drastically out of date. The team will require time to monitor current data and update their mental model.

Currently, CTs have several tools to help them in this updating process. One of these tools consists of the outgoing CT members who can provide information and insight into the current situation. These outgoing members are sophisticated and responsive but cannot help the incoming team for any extended period of time. Incoming teams also have text messages that provide information about events from the previous watch. Unfortunately, there is always a significant risk of errors of interpretation or omission with text messages.

The watch transfer procedure affects SA at all three levels (cf. Endsley, 1995). At level 1 SA, incoming team members must deal with a large volume of information in order to become familiar with the current situation. They have not only all of the normal tactical and procedural information associated with the watch but the messages providing information concerning previous events. This increases the risk of errors of perception and divided attention.

At level 2, incoming team members must rapidly update their understanding of the goals, ROEs, and other constraints associated with the mission. Until team members have done this, they will be unable to develop an accurate mental model of the situation. Even after identifying and incorporating changes to mission goals, team members must then attempt to re-integrate information with the updated goals to understand the situation.

Finally, at level 3, CT members must attempt to project events into the future based on their updated mental models. This projection will depend heavily on an understanding of previous trends and patterns – trends and patterns of which the incoming CT will have only secondhand knowledge. Being off watch creates large gaps in a CT's knowledge of events, perhaps disguising or obscuring important dynamic properties of events that occurred on previous watches.

These considerations indicate several broad needs for support of the watch transfer. CTs need some means to accommodate the large mass of data presented at the beginning of the watch. In particular, they need some tool that helps them summarize that data and integrate it in a mental model of the situation. Thus, any summary or description of events of the previous watch must be concise and clearly related to operational concerns. In addition, CTs require a tool to help update their mental models of mission goals and constraints.

To help CTs maintain continuity in their SA, they need a tool that provides a history of data streams across watches. This history must be categorized and abbreviated in some fashion to prevent CTs from being overloaded with data. In particular, the history should focus on important changes in pertinent dimensions of the situation. Effort should be spent on

determining the specific information needs of CTs and designing a system to meet those needs, which permits users to access and use information according to current needs.

4.3.2 Decision Making

The plans generated during mission preparation can pose a challenge for decision making. A large number of contingencies will have been addressed but the CT must determine the relevancy of each when coming on watch. Thus, the CT must determine how to prioritize pre-plans for the way ahead.

In part, this will depend on establishing good SA, especially level 3 SA (projection of events into the future). With awareness of how situations are likely to evolve in the future, the CT can identify relevant pre-plans and be prepared to implement them, perhaps keeping them highlighted in some way. Tools to support SA will have the added benefit of supporting decision making.

In addition to building good SA, however, the CT must explicitly link pre-plans to situational contingencies. As noted earlier, a potential problem of planning is that decision makers fail to see the connection between the current situation and pre-plans. Watch transfer offers an opportunity to make and strengthen these connections.

4.3.3 Teamwork

Watch transfer can be considered a weak link in the information transfer within teams and especially between teams. The discussion of SA indicates the threats to teamwork and communication within teams. At the onset of the watch, team members must work to update their mental models of the situation. As they do this, team members can develop differences in their understanding, disrupting team SA. Watch transfer will require, in addition to individual SA-building, coordination activities aimed at updating team SA.

Watch transfer also has profound implications for the interaction of teams across vessels. When a ship changes watch, there is a sudden shift of CT members. Thus, the entire make-up and character of the CT is radically altered. Other CTs in a task group who had been working with the outgoing team will now need to re-establish common understanding with the incoming team. If all ships in a taskgroup change watch simultaneously, this problem is compounded.

All this implies that CT can have varying degrees and qualities of team SA, depending on the point in time at which we look at the team. Teams will likely start off with relatively poor SA as team members catch up on events of the previous watch and coordinate their knowledge within the team. Teams will also have to coordinate across multiple teams within a task group, compounding the complexity of coming on watch.

When considering interactions of CTs, it is worth noting that watch transfer can alter the nature of CTs. Different teams will have different understandings of the situation and different approaches to decision making. When one team changes watch, other teams will have to accommodate the unique aspects of the incoming team. This could potentially affect the continuity of response of a vessel in its interaction with others.

4.3.4 Interface Design

The ecological approach to interface design focuses on describing the characteristics of a work domain. Thus, designers of C2 and decision support systems should model all of the important features of the work environment in which CTs operate. Watch transfer is one of those aspects but it is unclear that it has been systematically considered in the design of C2 systems.

We have already discussed several aspects of the watch transfer, the SA demands, and effects on decision making and teamwork. A more general concern is the way the interface supports CT members in their activities. The greater the time pressure and risk of the operational setting and the more demanding the information processing requirements, the more critical is continuity between modules of a command station. Because the cognitive demands on CT members will be very high, the interface must be designed to make movement between activities as easy as possible.

4.4 On Watch

Many of the issues discussed in relation to mission preparation and watch transfer apply to performance on watch. In particular, we have discussed the interaction of preparation and watch transfer activities on performance during the watch. This section will concentrate on issues specific to three broad activities of the watch:

- Surveillance.
- Threat Assessment.
- Threat Response.

4.4.1 Surveillance

Most of the CT's time on watch will be spent in surveillance activities. It is quite likely, except in times of active conflict, that the CT will not encounter any threatening situations. Nevertheless, the CT must remain vigilant to ensure rapid and appropriate response to any situation.

Thus, maintaining attention and SA are crucial to performance. What needs to be determined is how CTs react to prolonged, low-stress, uneventful periods. Essentially all research and theories of C2 and tactical decision making focus on threat assessment and response. These activities are clearly of paramount importance, which is why so much effort has been devoted to them. Unfortunately, this focus disguises the importance of surveillance and maintaining readiness when there are no immediate threats present or surveillance of a high traffic area.

There is reason to be concerned that these relatively uneventful periods could reduce SA and attention. An analogy can be drawn to driving an automobile on the highway. Although there may be numerous other vehicles on the road, a driver can easily be lulled into a state of reduced awareness when not confronted with road conditions, actions of other motorists, or other events that require meaningful processing in terms of the driver's goals (Gugerty, 1997). Research is needed to assess the impact of relatively uneventful watches on vigilance and readiness. Research in domains such as driving and air traffic control should provide useful insights into this issue.

4.4.2 Decision Making

As discussed earlier, theories of tactical decision making have typically focused on threat assessment and threat response, with little thought paid to mission preparation. As a result, intuitive theories have gained acceptance. These theories readily apply to time-stressed situations and can account for decision making under uncertainty.

Taking a broader view, however, has revealed the need to support both analytic and intuitive decision making strategies. Furthermore, it has revealed the need to support the synthesis of these strategies so that pre-plans developed during mission preparation can be used effectively during surveillance, threat assessment, and threat response. Thus, the primary decision support needs during these phases appear to be:

- Maintaining SA.
- Matching the current situation to previously established schemata (training, experience, planning).
- Selecting an acceptable response based on the retrieved schema.

One issue that has not been explicitly explored in the literature is the precise nature of decision making by the CT. We can distinguish two kinds of decision making relevant to CTs. The first is *team level* decision making, which is concerned with generating COAs on behalf of the entire team. This is the kind of decision making typically addressed by theories of decision making. The issue underlying team level decision making concerns what actions or responses the ship should make. An example of team level decision making is determining whether to engage a potential air threat.

The second kind is *individual level* decision making, which is concerned with determining the appropriate actions of individual CT members. The issue underlying this kind of decision making is determining the appropriate actions to take within the context of the CT to ensure that each member carries out his or her responsibilities and helps the CT function appropriately. An example is deciding whether to pass certain information to the ORO, based on knowledge of the situation, the ORO's role and workload, and the objectives of the team. Interestingly, individual level decision making has typically been investigated within the context of teamwork.

Nevertheless, it may be worthwhile to bring this distinction into discussion of decision making activities of CTs on watch. The CT will certainly be called upon to make important team level decisions during threat assessment and threat response. These decisions will be critical to the survival of the ship or taskgroup and the accomplishment of the CT's mission. During surveillance the CT may not be required to make many team level decisions. They will, however, constantly be making individual level decisions as part of the necessary activity to keep the team functioning.

It is wrong to view individual level decisions as unimportant. They may not have the immediate relevance of a team level decision about how to respond to a threat but they can have long-term, pervasive implications. Decisions concerning what information to communicate to other team members, the form of communication, and so on will affect team SA and, ultimately, team decision making. A decision to withhold information from the ORO because it seems irrelevant or because the ORO seems overloaded can turn out to be critical later on when the ORO needs that information to contribute positively to a team decision during threat response.

4.4.3 Situation Awareness

Perhaps the primary concern for maintaining SA during threat assessment and threat response is the sheer amount of information that must be processed in a very short period of time. The CT will receive a large amount of sensor data concerning potential threats, consult reference material to aid in classifying and identifying potential threats, review pre-plans, and coordinate resources across the task group in which it is operating. Unfortunately, the team may have only a few minutes or less in which to do this. Consequently, level 1 SA, the ability to take in information, is at severe risk in this phase and CTs need support for attentional and memory processes used in processing information.

In addition, we have already discussed the difficulties of creating good SA when the CT first comes on watch. This issue, of course, extends throughout the watch as the CT needs to be aware of what is happening at all times. In particular, the CT needs support to ensure that incoming information can always be related to command goals. Thus tools that simultaneously present and relate sensor data and goal-related knowledge (e.g., ROEs, mission objectives, etc.) will help the CT see the current situation in light of its goals.

4.4.4 Teamwork

Pigeau and McCann's (1998) analysis of C2 suggests that CT effectiveness rests, or should rest, on the establishment of common intent. That is, emphasis should be placed on conveying, explicitly or implicitly, to every team member the mission objectives. The basis for this claim is that team members are experts and can coordinate their activities flexibly to achieve mission objectives in the face of unpredictable operational conditions. Currently, there are no support mechanisms for common intent. Further research is needed to examine how CTs develop common intent within and between ships and what tools could assist CTs in doing this.

Stress is another issue that is very important to teamwork during threat situations. Although, this topic was not addressed in the current literature review, it has been extensively studied and should be considered in plans for the CPF upgrade.

4.4.5 Expertise

Expertise is typically considered in the context of an individual. Theories describe how a person acquires expertise, organizes domain knowledge, and expresses expertise. An area that has been neglected, however, is expertise in a team setting, where a number of experts in the same domain come together and act in concert. This issue could prove very important because experts in a team must coordinate their individual, and to some extent unique, expertise.

Consider two people who have served in the Navy. It is unlikely that they have followed the same career path, served the same length of time, or had exactly the same training experiences. Because expertise arises so directly from direct experience and practice, these two people will not have exactly the same knowledge bases. Nor will they have organized their knowledge around exactly the same principles. These differences are likely to be greater the more divergent their career paths have been. An ORO who served previously as a SWC may have a very different understanding of C2 and CT functioning than an ORO who previously served in communications or as a submariner.

Differences in expertise could have a number of negative effects on team level decision making. Such differences could make communication within the team more difficult, impair team SA, or weaken common intent (cf. Pigeau & McCann, 1998). Research is needed to explore not only the formation of expertise but how it is used in teams to achieve team decision making.

5. Methodology

This section reviews methodological issues and established methodologies for studying issues related to naval C2 and decision support. The term methodology can apply to at least two broad areas. First, there is methodology pertaining to the design, implementation, and evaluation of C2 and DSS systems. In this case, methodology refers to a range of engineering and management. Second, there is methodology pertaining to scientific inquiry. In this case, methodology refers to empirical and analytic techniques for making observations, testing hypotheses, and gathering data needed to advance our theoretical understanding. We examine methodology in both senses in this literature survey. A great deal has been written on the design process for naval systems and this work has clear relevance to the upgrade of the HALIFAX class. These design methods, however, tend to rely on user-centered, iterative approaches in which prototype systems are continuously evaluated in relation to the human element as well as the technological. Consequently, design methods must be supported by empirical techniques that can provide high quality, reliable data. In addition, our understanding of C2, human decision making, and many other domains is still very limited. By reviewing empirical methods, we can expand our ability to test current theories and promote new ones.

5.1 C2 and Decision Support Design Process

A design process is an overall approach to identifying what a DSS should do, how it should do it, and how the system can be built to perform its function. This process guides not just the actual building of a DSS but the conceptual analysis of the tasks and problems to be solved and the evaluation strategy.

A design method should identify issues to be resolved in three areas (Vineberg & Bryam, 1994):

- Organizational: control of the development process from concept design to prototyping to testing.
- Functional: specification of user functions, system functions, command system to support user functions.
- Physical: identifying requirements for hardware and software.

For purposes of this review, interest focuses on the functional design issues. A design method for naval tactical C2 decision support must address the kinds of tasks and missions faced by the commander and CT. Recall that Anderson (1990; Section 2.1.2) has identified four main conceptual areas:

- Information management.
- Platform management.
- Resources management.
- Tactical management.

Each of these represents a broad component of the C2 task and each presents problems for effective decision support. Information management can be difficult because of the volume of information and its ambiguity. Platform management is an integral part of tactical operations and the C2 team needs support to manage the ship. In addition, the commander must be concerned with managing the ship's limited resources (ammunition, fuel, etc.). Finally, the tactical situation will be the prime focus of the C2 team during critical events.

In addition to these four areas, we should add team management. As discussed previously (Section 2.5), teams require coordination to function. This coordination may be implicit or explicit but team members engage in some teamwork activities to maintain the team. Thus, another aspect of C2 concerns issues related to communications among team members.

These conceptual areas do not dictate a particular methodology but indicate the issues that must be addressed by a methodology.

5.1.1 Principles of Design

A consensus appears to be developing as to the general design approach that should be applied to naval C2 systems. This approach is not specific to the design of DSSs but can be directly applied in this domain. The following principles indicate a desire for a method that is user-centred and flexible to accommodate useful input to the design of the system and respond to changes in needs or procurement constraints.

5.1.1.1 User-centred Design

The central goal of decision support is to present decision makers with tools that improve their decision making abilities. Many researchers have argued that the best way to do this is to design tools that parallel the cognitive processes of expert human decision makers (e.g., Hutchins, 1996). Thus, the design process must have as its first principle to gain an understanding of how the human performs the tasks. Even if one adopts an analytic or expert system approach to decision support, success depends on the operator being able to make use of the system's output. As a result, many researchers advocate a user-centred design approach in which an analysis of how people perform is the focal point for developing design concepts (e.g., Adelman, 1992; Robson, 1997; Schuffel, 1994; Vineberg & Bryam, 1994).

User-centred design begins with a definition of user requirements (Adelman, 1992). These are the information, cues, aids, and procedures that the human operator uses to perform tasks. They serve as the elements that must be incorporated in the design of the system.

The next step is an analysis of how the user employs these elements. The designer must distinguish which functions should be performed by the human and which by the machine (Schuffel, 1994; Robson, 1997). The division of tasks between the operator and the system will be a key determinant of operator workload and, hence, effectiveness. This process depends on a formal task analysis to decompose the overall job into its component procedures and steps. When the human's tasks have been specified, the designer must then create a framework identifying user functions and goals.

This approach emphasizes the cognitive processes of the user. The value of doing this is that the increasingly complex systems being deployed in the Canadian Navy place human operators in increasingly supervisory roles (Vineberg & Bryam, 1994). Operators engage in more high-level decision making, monitoring, and troubleshooting and less in manual or computational tasks. Consequently, the features that will define a user-friendly system arise increasingly from the cognitive level.

5.1.1.2 Top-Down Versus Bottom-Up Approaches

User-centred design embraces both *top-down* and *bottom-up* design processes (Cochrane & Foley, 1991). That is, it employs input from both designers and users.

The top-down approach is conceptually driven. It begins with some definition of system effectiveness in terms of the desired outcomes of deploying the system (Leroy & Hoffman, 1996). This functional analysis refers to theoretical models of the command process and a formal analysis of the C2 problem space. The aim is to develop a process model of how the system will function and achieve its design purposes. The main purpose of C2 systems is to increase the efficiency of commanders and CTs in organizing and directing fighting forces. Thus, system effectiveness is measured either in terms of C2 performance (throughput, error rate, etc.) or effectiveness (how well the system functions within the operational environment, probability of detection, reaction time, etc.) (Leroy & Hofmann, 1996).

The top-down approach employs analytic techniques and describes the formal problem structure that the system works within. This approach, however, is somewhat removed from the real world and depends on theoretically derived concepts and measures. Top-down design can never capture the complexity of real-world situations or all the factors affecting C2.

The bottom-up approach is observation-driven. It begins with an analysis of how experts perform their functions in the field. This analysis can refer to many sources, such as doctrine, interviews with staff officers and active personnel, task analyses, and so on (Arnfield & Smith, 1996). The goals are to identify what would make an effective system based on the views of the people who will actually use the system and to avoid problems already identified for earlier systems. The assumption is that these SMEs will be able to identify design concepts.

Technical Services Attributes	Applications Attributes
Availability (system must be operational at all times)	Interoperability (usable with other systems, current and future)
Survivability (withstand attack and/or loss of some components)	Confidentiality (restrict access to data)
Robustness (adaptable to environmental conditions)	Integrity (preserve completeness and accuracy of data)
Maintainability (ease of repair or replacement)	Customizability (adaptability to new functions)
Computation capacity	Quantity of Information (provision of all needed data)
Mobility	Bandwidth (match rate of data presentation to user's capabilities)
Portability (usable in new platforms without extensive modification)	

Table 5.1 – Technical Effectiveness Criteria for C2 and Decision Support Systems (from Leroy & Hoffman, 1996)

Bottom-up design considers three classes of effectiveness criteria, as described by Leroy and Hoffman (1996). The first consist of technical services attributes, such as hardware and software attributes that will affect how the system performs. SMEs who maintain systems in the field are asked to identify requirements in a number of areas (see Table 5.1). Overall, these criteria determine how the physical system will be deployed, maintained, and used. SME input helps the designers identify real-world concerns about the system and set requirements and standards that must be met before the system is ever built.

The second set of criteria deal with attributes determining the effectiveness of the user's work. SMEs who will be the chief users of the system identify design elements that can facilitate or impair the performance of tasks with the system. These are broken down into two subsets, criteria related to information quality and criteria related to user motivation in using the system. Table 5.2 lists the attributes in each subset. Overall, these attributes deal with the user's interactions with the system and how well the system would support the user in performing tasks. The goal is to develop specific criteria regarding information access, speed, and so on, that can be translated into design standards.

Attributes Related to Information Quality	Attributes Related to User Motivation
Selectivity (ensure relevance of data)	User Acknowledgement (confirmation of inputs and communications)
Accuracy (maximize accuracy)	Knowledge of Consequences (feedback about potential outcomes of user decisions)
Comprehension (facilitate understanding)	Knowledge of Common Goals (feedback about team task)
Response time (react within established criteria)	Confidence in Automated Processes
Timeliness (present information when it is needed)	User-Friendly (attractive, adaptable interfaces usable by different users in various states such as stress)
Ease of use (access and control of data)	Participation in Design Process (user involvement in setting design goals and concepts)
Training time (minimize)	
Decision response time (work within user's time constraints)	
Uncertainty management (deal with ambiguous data)	
Simulation (test COAs before implementing them)	
Information Age Management (store large amounts of data in organized, accessible format)	

Table 5.2 – User Effectiveness Criteria for C2 and Decision Support Systems
(from Leroy & Hoffman, 1996)

Finally, there are ineffectiveness criteria that indicate problems that can arise in poorly designed systems. These represent unintended consequences that must be specifically dealt with in the design process:

- System isolates the user from the team (Loneliness).
- System facilitates inappropriate inputs from others (Interference).
- System induces a strategy of waiting for a complete tactical picture (Wait and See).
- System fosters over-confidence in recommendations of the system (Judgment).
- System distances the user from the real situation so that the user forgets the system is not a game (Wargame Mindset).

Although the bottom-up approach identifies real-world issues it can be limited by its reliance on SMEs (Leroy & Hofmann, 1996). SMEs have direct experience rather than theoretical insight into the task domain. In assessing system requirements, SMEs may think only in terms of what currently exists. Users often have no idea of the capabilities of new technologies and are unable to address how new technology might be used. Even worse, SMEs may be limited in how they conceive of tasks, considering functions only in terms of how things are currently done. One of the primary benefits of a conceptual analysis of the task domain is that it can identify system elements that might enable new, more effective, strategies for performing tasks.

5.1.1.3 Rapid Prototyping

Previously, design methods followed a *waterfall model* in which development followed a strongly linear path (Anderson, 1990). The process began with analysis of the system requirements, followed by functional specification, architectural design, then implementation, in that order. It was only after implementation that users would be consulted again, the system evaluated, a new cycle of development begun to revise and correct any flaws in the system.

The problem with this method is that there is a long user feedback loop. After an initial survey of users, the design process proceeds without any user feedback until a prototype system has been built. Thus, designers invest a great deal of resources in a system that may not meet user needs very well (Cochrane & Foley, 1991).

In response to this problem, designers in the United States and elsewhere have adopted a process of *rapid prototyping* (Anderson, 1990; Hutchins, personal communication, 1998; Manning, 1991). This is a method of "build a little, test a little" (Leroy & Hofmann, 1996). It uses a tighter user feedback loop to avoid large wasted effort and iterates the development cycle more quickly and more frequently to respond to feedback and changes in Navy doctrine and missions (Anderson, 1990).

The purpose of rapid prototyping is to develop early on a comprehensive representation of the overall design so that all system elements and their interrelations can be evaluated (Matthews et al., 1997). The representation must be economical to minimize costs and delays but comprehensive enough to allow designers and potential users to comment on the system.

The basic method follows the steps of the waterfall method but gathers user input at each step. Thus, users are consulted for requirements specification, functional analysis,

architectural design, and implementation. In addition, rather than attempting to generate an actual system that can be implemented in the early stages, the method focuses on developing prototypes and describing levels of functionality. A prototype is a shell of the system that provides most of the interface functions and can simulate what the system is supposed to do (Manning, 1991). The prototype should illustrate the system's functional capabilities to the level of detail needed for the evaluation of the system. It is a tool for refining system requirements and the precise nature of the prototype will be determined by the purpose for which it was constructed. Users are brought into the design method to validate the prototype at each stage, commenting on the function and design exhibited by the prototype. The prototype can then be quickly and easily refined on the basis of this input.

Rapid prototyping is valuable because it reduces costs while allowing more data to be collected from users. In addition, it allows the designers to be more responsive to user input because they have not invested large amounts of their limited resources into a working system. Instead, the designers can radically change a prototype at any stage in the design process to respond to unanticipated requirements or constraints.

Matthews et al. (1997) list several important features of rapid prototyping:

- Rapid prototyping is a recurrent, iterative process that requires constant evaluation and redesign.
- A prototype is a temporary representation of how the system will work and must be updated as new insights are made.
- The focus of a prototype can be on all or part of the system.

A common way of implementing a prototype is through a *storyboard* (Matthews et al., 1997). A storyboard captures a system in a series of "pictures" that form a "narrative of navigation" through the functionality of the system. The pictures can consist of artist sketches, computer screen displays, or interactive computer programs. The purpose of the storyboards is to illustrate the functionality of the user interface to the system so that potential users can experience how the system will work. In particular, input and output screens are sequenced to emulate the sequence of interface actions as the user interacts with the interface (Miller et al., 1992). At higher levels, prototypes can also be implemented by part-system test beds and full-scale simulations (Matthews et al., 1997). The advantage of storyboarding, however, is that it is low cost and rapid, making it easy for designers to test interface concepts and features and revise them based on user feedback.

5.1.1.4 Potential Problems in the Design Process

A number of problems can arise in the system design process. Some are due to institutional actions or inaction. Others are due to constraints of technology or the task domain.

Cochrane and Foley (1991) noted that a serious problem in the development of a tactical C2 system for the US Marine Corps was the lack of consensus concerning the centralization versus decentralization of C2 functions. Without a clear doctrine, the designers were unable to develop stable requirements for the system. Without clear requirements, the system as it developed failed to satisfy either potential users or senior Marine Corps officers.

The first step of the design method must be to establish agreed upon principles. Unfortunately, even if there is agreement, the requirements developed may still be inadequate. Naval C2 is a complex domain requiring a great deal of team interaction. It can be difficult to capture the effects of team interaction even when surveying SMEs (Miller & Woods, 1997). Analyses of user needs should be conducted with multiple scenarios in mind.

Finally, the development process can be limited by available resources of time, money, and personnel (Anderson, 1990). Such limits can force trade-offs that ultimately reduce system effectiveness.

5.1.2 Development Methods

5.1.2.1 A Rapid Prototyping Methodology

Several researchers have used the above principles to developed detailed system design methodologies. Based on the methods proposed by Adelman (1992), Cochrane and Foley (1991), and Schuffel (1994), we have created a composite rapid prototyping methodology. This method combines elements from all three sources to provide a step-by-step outline, presented in Figure 5.1 below. Items in boxes represent fairly discrete steps in the overall process. Steps are organized sequentially to indicate the overall flow of the process. Curved lines represent steps that can be iterated as many times as needed during the process. Steps linked together on a horizontal line are highly related, iterative sub-processes.

The method follows three broad stages. The first is *functional analysis*, in which the design team describes the specific requirements and goals for the system (i.e. what the system is supposed to do) and determines whether development of a system is feasible. In the feasibility assessment, designers determine whether the proposed system could meet the following criteria (Cochrane & Foley, 1991):

- Compatible with current doctrine and established practices.
- Technically possible.
- Limited complexity.
- Can be procured effectively.

The second stage is rapid prototyping. This stage involves the identification of system requirements based on user needs. Requirements are a joint function of the task and the users, who will have preferred strategies for performing the task. At this stage, designers select the specific analytic methods to create the functional model of the task. Designers use these methods to create the functional model that describes what users do to solve the task and the information they need.

The third stage is *implementation*, in which software and hardware issues are resolved and a working model is developed. The system is then deployed and tested.

The rapid prototyping methodology is a user-centred approach and was a reaction to earlier technology-centred methodologies. A potential problem, however, is that this method can overlook important technology issues until the implementation stage. The feasibility assessment would certainly contain an analysis of technological capabilities and requirements but this could only examine projections of technologies expected to be available at the time of completion. In this method, it is possible that designers may

find that technology is not sufficient to implement a system as designed. Furthermore, by focusing on user requirements and feedback, designers run the risk of neglecting functions and design features made possible by new technologies.

5.1.2.2 User-Centred Design Model

Webb et al. (1993) have described another rapid prototyping methodology. This methodology seeks to involve HF and user input at every stage of development from the initial statement of requirements from the user's perspective to the final field testing of the system with typical users. Figure 5.2 below diagrams the design cycle with potential HF methods superimposed.

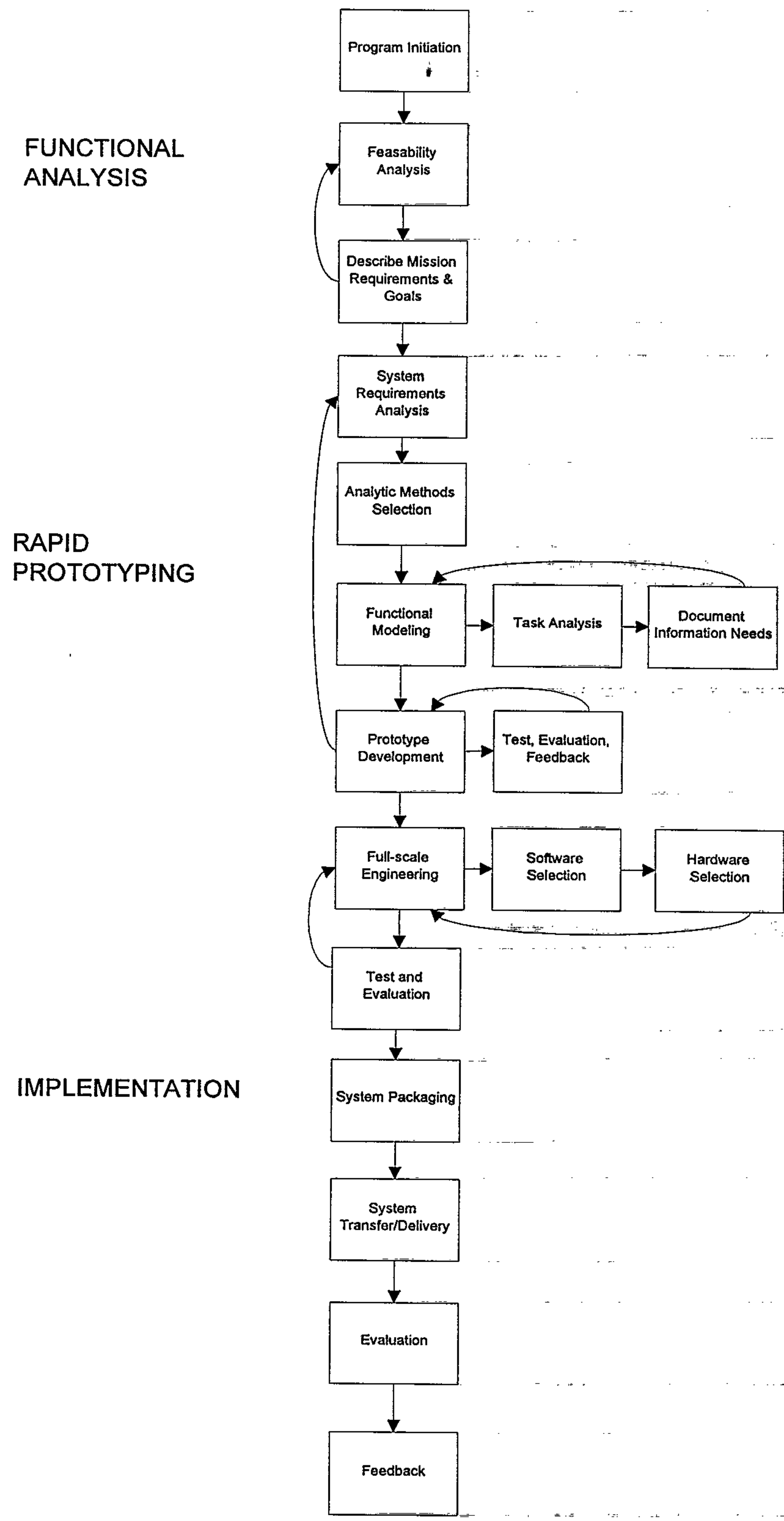

Figure 5.1 – Rapid Prototyping Methodology
(From Adelman, 1992; Cochrane & Foley, 1991; Schuffel, 1994)

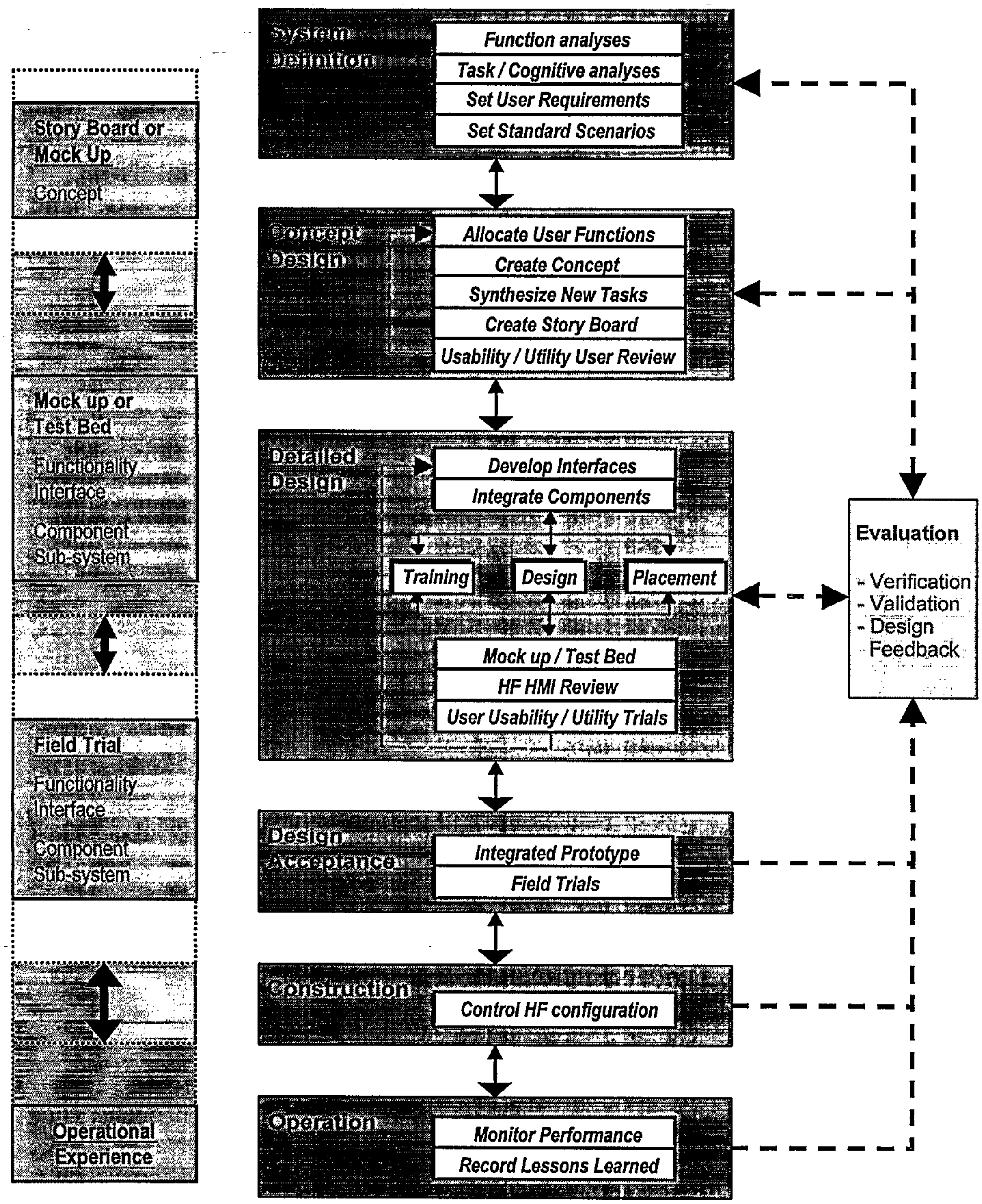

Figure 5.2 – User-Centred Design Model (from Webb et al., 1993)

The user-centred design method consists of six basic steps.

System Definition. This stage consists of the definition of system requirements and an analysis of system functions. The *system requirements analysis* examines the HF implications of each part of the system and determines the criteria for eventual testing of the system. The *function allocation* uses numerous sources (e.g., analyses of similar systems, doctrine, SMEs, training manuals, etc.) to describe what the system is to do.

Preliminary Design. This stage examines the feasibility of different ways of implementing the functional requirements of the system. Function allocation allows designers to decide which functions the system will perform and which the user will perform, taking into account hardware, software, and human performance factors. During task synthesis designers organize functions into coherent tasks and make sure these tasks are consistent with HF criteria.

Detailed Design. This stage creates a concrete working design of the system. Designers create a *task description* by expanding the task synthesis. The task description describes, in detail, task components and activities in terms of a number of factors (e.g., time lines, information requirements, skill requirements, etc.). Designers conduct a *task analysis* to evaluate the demands placed on the user to complete each function. They also consider training, interface, and user placement issues. The end result is a *prototype* of the system.

Validation. In this stage, designers establish operational criteria based on earlier analyses and test the prototype. The goal of testing is to determine that the combination of system functions and personnel and equipment assumptions can accomplish the operational requirements.

Construction. In this stage, a fully operational version of the validated system is built. It is important that the construction process allow modifications to be made but only after review by the design team to ensure compatibility with the original design goals.

Operation and Monitoring. When the system is in use, designers must monitor user performance and acceptance and gather feedback to determine how well the system is functioning.

5.1.2.3 Interface Design

One significant component of system design is the design of the user interface. The software development industry can be a great asset in this area because private companies, such as IBM, Microsoft, and Bellcore have developed methods for designing and evaluating interfaces. Furthermore, many of these companies make their interface guidelines available, free of charge, on the World-Wide Web (WWW).

The specific methods are too numerous to review here (see Matthews et al., 1997). In general, these methods parallel the rapid prototyping method discussed above. For example, one site on the WWW presented a Graphical User Interface (GUI) design methodology that emphasized user-centred design and iterative design-feedback loops (http://www.dordt.edu:457/VTCLG/vtclgD.style_standards.html). The method proposed four basic steps:

1. **Create design specifications**: define the objectives and features of the system using input from actual users who understand the task, intentions, strategies, and decision processes involved.

2. **Use prototypes**: create models that functionally illustrate design features and test their value with users, eliciting specific feedback and comments on the functionality of the design.
3. **Assess and test**: starting early in the process, conduct usability assessments (e.g., walkthroughs, focus groups, heuristic evaluations) and have users generate success/failure data as well as lists of needs and wants.
4. **Iterate**: as each evaluation indicates problems, revise the prototype and test again, progressively refining the design to meet user needs.

5.1.3 Recommendations

- Expand consideration of context in the design process to address questions of what situations the CT will operate in and how these situations will affect design assumptions.
- Conduct further research to develop a formal model of the C2 domain specific to the CPF (roles and missions, tasks and functions, problems faced, ROEs, optimal ways to solve problems, etc.).
- Establish clear and consistent goals and principles for a system at the outset.
- Identify information management, platform management, resource management, and tactical management issues.
- Adopt an iterative, user-centred approach to DSS design.

5.2 Research Tools

This section reviews experimental methods, analysis techniques, simulations, and other resources that can be used to conduct tests and evaluations in the design process.

5.2.1 Laboratory Simulations

Among the important research tools are research laboratories. This is especially true for the study of naval C2 and decision support because of the technical complexity of the domain. To study how decision makers perform aboard ships, researchers need facilities that can simulate the complexity of the tactical situation, if not the actual equipment and resources of the OR.

Most facilities for studying tactical decision making serve to conduct simulations. Consequently, most research concentrates on assessing performance of experienced personnel in realistic situations, performing real-world tasks (although some research examines the performance of novices in contrast to experts). This approach stands in contrast to a great deal of research in the behavioural sciences, such as psychology, which focus on performance of specific, theoretically-derived tasks. The difference, in part, reflects the military's concern with predicting actual performance in the field, whereas behavioural science wishes to test hypotheses regarding theories of behaviour. The difference also reflects the tacit belief that the military tactical decision making domain is so specialized that assessments of performance in abstract or non-tactical domains will provide limited insight into theories of tactical decision making.

5.2.1.1 Scenario-based Versus Cellular Automata Simulations

Existing simulation facilities can be divided into two contrasting approaches. One is *scenario-based.* Here, the goal is to create a detailed simulation of a particular set of

events. The simulation sets up a script that describes the tactical situation, the units involved, and, most importantly, the specific actions that will be taken by the units. The simulation describes the sequence of these actions and typically contains very few points at which the actions of the participant will dramatically affect the events in the scenario. For example, in an anti-air simulation, participants may monitor a track and attempt to identify it and its intent. Participants' efforts to gather information and take defensive steps do not alter the simulated track. Thus, participants respond to a preconceived idea of what a potential air threat might do rather than to an intelligent agent in the simulation.

The scenario-based simulation proceeds as the participants are presented with sensor data corresponding to the events in the scenario. Typically, their view is through realistic interfaces and participants experience events as they would in actual tactical situations using actual equipment. Participants respond and also act through realistic interfaces as well.

An example of a scenario-based simulation is the Aircrew Evaluation Sustained Operations Performance (AESOP) facility (Elliot et al., 1996). This simulation is designed to study decision making among team members in the AWACS aircraft. Participants play roles in plausible simulated scenarios with high threat, time pressure, and ambiguous data. One feature of AESOP that sets it apart from many other scenario-based simulations is that it makes use of authentic equipment and displays. Thus, it conveys a greater sense of realism than other simulations. In addition, experienced personnel serve in various roles in the simulation to provide participants with more realistic interactions with entities outside the team.

The benefits of the scenario-based approach include (Matthews et al., 1997):

- Operational relevance.
- High diagnosticity of system functions.
- Systematic capture of data related to users underlying information needs.
- Standardization for comparison purposes across successive prototypes or comparable products.
- Results experienced in terms relevant for design improvement.
- Results expressed in terms understood by the user community.

The other approach to simulation is to use the computational power of computers and create virtual entities in a tactical situation. The advent of low-cost, powerful computers has provided the technology necessary to simulate more than specific scenarios but the entire tactical world in which people can operate. This is the *cellular automata* approach, which creates a simulated world with numerous entities and allows participants to interact freely with those entities. The goal is to create a low-cost simulation environment in which an unlimited number of events can unfold. Thus, the simulation does not contain detailed scenarios. Instead, it specifies the forces, terrain, and other conditions of the situation, then allows the participant to interact with those elements. The actual events that occur are determined by the actions of the participant and the rules governing how simulated units act and react to events in the simulation.

A cellular automata simulation works by creating data structures corresponding to various entities in the simulated world, such as aircraft, ships, missiles, and so (e.g., Federico et al., 1991). Programmers specify comprehensive rules that govern how each entity acts and interacts in the simulated world; e.g., movement, initiating attacks,

defensive actions, etc. The rules indicate the conditions necessary to initiate some action, how the action is implemented, the results of the action, and the other factors such as timing and sequencing. The program does not specify a particular series of actions that the entities will take or any events that must occur. The simulation specifies only the initial state of the simulated world and the rules by which the entities themselves will determine what actions to take.

Thus, cellular automata simulations are highly flexible. They require a large investment of effort initially to model the world and devise comprehensive rules for the entities inhabiting the world. After completion, however, the start-up conditions can be easily and rapidly changed to create new problems for participants. Also, given a particular start-up configuration, the simulation will never run the same way twice because the simulation responds to every action of the participants. Even small differences in how one acts can have large and compounded effects on the actions of simulated entities throughout the course of the simulation. Thus, cellular automata simulations simulate field exercises rather than particular scenarios or situations.

One resource for cellular automata simulations is the computer game industry, which has created numerous simulation-based war games. Games like *Total Annihilation*, *Harpoon*, and *Steel Panthers* have powerful simulation engines and can simulate numerous tactical problems in naval, air, and land contexts. The games themselves are not necessarily directly applicable to studies of tactical decision making but the simulation software is. With some modification to the units, time scale of play, and so on these games could provide a training and experimental testbeds with simulated environments and units that closely correspond to real situations and forces. Researchers would have to change the databases of the gaming software but could take advantage of the sophisticated computer modeling techniques already developed. In fact, this approach has been used already to study decision making (Federico et al., 1991; Omodei, McLennan, & Wearing, 1998).

Cellular automata simulations generally have little fidelity to the combat situation, at least as a combatant would experience the situation. They use simple interfaces that provide tactical information in a form unlike the way personnel receive information in the field. However, if effort is devoted to designing units, environments, and tactical situations that are realistic, these simulations can be used to examine tactical decision making processes. In addition, interfaces that constrain how information is presented to the participant can increase the realism to the participant while maintaining the benefits of speed and flexibility.

An example of a quasi-cellular automata approach is the Canadian and US Army's JANUS (Hakala, 1995). This is an interactive, two-sided, closed, stochastic ground combat simulator. JANUS uses digitized images of actual terrain to establish the environment and model terrain effects such as movement, lines-of-sight, and so on. In a JANUS simulation, opponent forces are placed in some initial set-up and the participants' team given mission objectives. The participants use low-fidelity simulations of battlefield equipment to direct their forces in the simulated environment. JANUS employs human SMEs to control enemy forces. Thus, humans take the role of the rules that govern unit actions in a true cellular automata simulation. The human operators, however, can interact in many ways with the participants and capture the goal of the cellular automata approach, which is to simulate a tactical environment.

Currently, scenario-based simulations are much more common than cellular automata simulations. The following sections briefly discuss some existing laboratory simulations beginning with several scenario-based simulations then a cellular automata simulation.

5.2.1.2 Decision Making Evaluation Facility for Tactical Teams (DEFTT) Laboratory

DEFTT is an unclassified, low-fidelity simulation testbed developed for the US Tactical Decision Making Under Stress (TADMUS) program (see Hutchins & Kowalski, 1993; Radtke & Frey, 1996). It simulates five computer workstation positions (hardware and software) used on the US Navy's AEGIS Class cruiser. The workstations are controlled by a dedicated server that can run complex anti-air scenarios that contain realistic sensor data.

The network stations allow researchers to study team performance as well as perform training exercises. The workstations simulate the Command and Decision (C&D) console assigned to the CO or Tactical Action Officer (TAO), and the Tactical Display Stations (TDS) assigned to the Electronic Warfare Supervisor (EWS), the Anti-Air Warfare Coordinator (AAWC), the Tactical Information Coordinator (TIC), and the Identification Supervisor (IDS). Scenarios present all the data and functionality necessary for the team to detect, identify, and engage air and subsurface tracks. Each station records and time-stamps the keystroke and trackball actions of the user. All communications are recorded as well. Thus, DEFTT can capture a great deal of detailed data on individual and team performance.

Scenarios developed for DEFTT seem to emphasize critical situations and not planning and preparation. There seems, however, to be no reason why these activities could not be included. A more serious restriction is that DEFTT simulates only the shipboard environment; there are no elements to simulate task group or other groupings of vessels. Thus, DEFTT lacks a degree of complexity of real operational settings.

DEFTT simulates the world only as members of the OR team would perceive it. That is, there is no overview of the tactical situation and the only data available are the simulated sensor data. This facilitates study of decision making under realistic tactical situations but limits the possibility to study potential new systems that might present greater information about the situation. DEFTT can also be used as a training tool to facilitate practice of procedural skills and learning of cues to recognize tactical situations. It does not, however, provide a means for teaching higher-level tactical concepts.

Also, as a training tool, DEFFT does not meet many requirements of a multimedia instructional tool (see Section 2.4.3.4). DEFTT does provide a logical organization of content in terms of missions and resolving tactical problems and it requires active participation by the learner (Hutchins & Kowalski, 1993). DEFTT, however, does not explicitly guide learning toward an objective. Thus, it is more of a practice tool to reinforce concepts already learned than a training tool for learning new material.

5.2.1.3 Georgia Tech Anti-Air Warfare Coordinator Simulation Suite (GT-AAWC)

The GT-AAWC is an enhanced version of the TDS workstation of DEFTT (Radtke & Frey, 1996). GT-AAWC creates synthetic team members for the AAWC to interact

with. It also presents very detailed data concerning air tracks. Thus, GT-AAWC allows more detailed study of the AAWC position by providing greater functional fidelity.

This simulator collects data in the same fashion as DEFTT and can be used to conduct empirical research or provide training. As an empirical tool, GT-AAWC is suited only to the study of individual decision making and performance because participants interact with synthetic teammates that are based on researchers' understanding of the OR and may not support realistic interactions with the AAWC. As a training tool, GT-AAWC can be used as a part-task trainer to demonstrate and practice components of the AAWC's responsibilities in realistic scenarios.

Like DEFTT, GT-AAWC represents tactical situations from the perspective of the AAWC and provides no overview of the tactical situation. Unlike DEFTT, however, GT-AAWC does present additional cues designed specifically to help teach the participant how to recognize certain tactical situations. It also provides summary data about performance as feedback. GT-AAWC is designed to fulfil a training role and is consistent with several multimedia instruction techniques. The focus on feedback is a definite aid to learning. GT-AAWC also organizes the content logically in terms of realistic missions and it requires the learner to actively analyze the content of simulations. GT-AAWC provides some guidance in scenarios by highlighting important cues.

5.2.1.4 Team Model Trainer (TMT)

Radtke and Frey (1996) also describe the TMT, which serves more of a training than research function but which can be used as a tool in some empirical and analytic techniques. TMT presents two explicit options, training and simulation. The training option contains multimedia instruction in six domains (equipment, task, defence system, team, ship, and situation) at three levels (declarative, procedural, and strategic knowledge). The simulation option allows participants to apply their domain knowledge by observing or participating in event-based scenario simulations. In a simulation, participants observe an air defence scenario from the perspective of any of the five positions in DEFFT. Participants can take an active role by performing the tasks of a given station using a low-fidelity representation of the console display. There is an observe mode in which participants monitor video and audio presentations of a team performing the scenario. The observation mode could be useful for interview and walk-through techniques such as cognitive task analysis (see Section 5.4.1).

5.2.1.5 Tactical Naval Decision Making System (TANDEM)

TANDEM is a low fidelity simulation of C2 and communications in the Combat Information Centre (CIC) of US Navy combat vessels (Weaver et al., 1995). This simulation was specifically designed to study team performance and focuses on scenarios that require extensive sharing of information by team members. Scenarios focus on threat detection and identification. In a typical scenario the team is presented with numerous radar contacts and must detect potential threats and respond appropriately.

TANDEM was designed to investigate factors that could affect team performance, such as task interdependence, time pressure, task load, and ambiguity of data. It is implemented on highly configurable PC workstations that allow researchers to quickly

change the set-up to vary these and other factors. The workstations collect user input data and communications are recorded for later analysis.

TANDEM is very similar to DEFTT but sacrifices some degree of fidelity to the shipboard tactical environment for greater flexibility in conducting empirical investigations. Weaver et al. (1995) discuss several other low fidelity simulation testbeds that may be adaptable to naval C2 issues.

5.2.1.6 Battle Management Assessment System and Raid Originator Bogie Ingress (BATMAN & ROBIN)

BATMAN & ROBIN is one of the few examples of the cellular automata approach for naval C2 (Federico et al., 1991). In some respects it is similar to JANUS but is a completely computerized simulation system. BATMAN & ROBIN is designed as a PC-based performance measurement system that can also be used for training, development and evaluation of tactics, generating scenarios, and planning. The system makes use of high resolution graphics but has very low fidelity to the shipboard environment.

BATMAN & ROBIN models the naval tactical environment. Participants control a ship and its resources (including deployable air units, missiles, and other weapons). Entities in the environment include surface and subsurface vessels and aircraft.

The key to BATMAN & ROBIN is the AI program that controls units in the simulation. It uses the knowledge-based Finite State Automata (FSA) approach (see Federico et al., 1991, for discussion of rejected options). There are no pre-scripted scenarios that units follow. Instead, each unit is defined by a comprehensive database that describes its characteristics, *action states*, and rules that execute transitions between states.

The characteristics of a unit are the parameters on which it can vary. These include indicators of physical qualities, such as position and integrity or "health." Other parameters describe the unit's operational and behavioural characteristics, such as readiness, fuel levels, ammunition, offensive or defensive posture, attacking, withdrawing, and so on. Finally, each unit has functional characteristics that determine how it interacts with other units, including goals and intent.

An action state is the current status of the unit as indicated by the configuration of values of the parameters that describe the unit.

Transition rules indicate the conditions under which changes in parameters will occur. They are formal rules that can be implemented in a computer program and take the form of "if-then" productions that indicate a discrete change in the unit given certain well-defined conditions. For example, a transition rule for an aircraft might take the form (Federico et al., 1991):

Rule: (when true) Blue air kill necessary, (cause the transition) kill enemy

This rule states that when the conditions require the destruction of a blue air unit, then the unit will initiate its productions for killing the enemy. Thus, the rule would call upon other rules to evaluate, for example, the proximity and value of blue air units. If a valuable blue unit was in range, the rule would then call upon other transition rules to move the unit into a state of attack readiness. This example illustrates that the transition

rules must be nested, with rules referring to other rules to define their conditions and calling up other productions to enact more detailed actions.

The success of BATMAN & ROBIN depends on the comprehensiveness and accuracy of the production rules. The set of rules must be able to respond to all possible situations that could occur in the simulated world. Furthermore, to be of value as a research tool, they must reflect the actual behaviours of units in the real world. Thus, the development of the simulator, and any simulator like it, requires extensive analysis and effort.

Federico et al. (1991) list all of the transition rules of BATMAN & ROBIN and give a description of the kinds of situations it supports. As mentioned, the simulation is not faithful to shipboard conditions. It does not simulate equipment or conditions. It does not even simulate functional aspects such as the amount or quality of data available. Nevertheless, this kind of simulation warrants investigation as a research tool. Its ease of use and flexibility allow for rapid experimentation.

5.2.2 Test Options

There are several options for providing evaluation capabilities of the Canadian Navy (see Webb et al., 1993, for a review of evaluation options for the Canadian Army). These options include:

Dedicated Test Bed. This would consist of a permanent facility for the simulation of the OR of the CPF. The intent is to achieve high fidelity to shipboard OR, although practical and cost constraints may limit this. The Canadian Navy's training facility in Halifax could potentially serve in this capacity.

Low Fidelity Simulations. This option consists of a small room containing at least four computer workstations configured to simulate the back row positions (CO, ORO, SWC, ASWC) on board the CPF. Low fidelity simulations provide the capability to examine a wide range of OR functions and decision making but not interactions with OR equipment (Weaver et al, 1995).

Field Exercises. It is also possible to conduct studies on board active CPFs. In this case, scenarios can be presented on actual OR equipment to provide the highest possible fidelity to actual tactical situations.

Cellular Automata Simulations. This approach consists of low fidelity simulations that emphasize high level tactical concepts rather than practical training. Cellular automata simulations provide a means to study decision making and related factors in a fast and economical way. Because of their speed, these simulations allow decision makers to attempt many approaches to solving problems and immediately see the consequences of their actions.

Laboratory Studies. Traditional experimental techniques provide controlled studies of theoretically relevant issues and applied problems. The drawback of these techniques is that they are expensive and time consuming, requiring extensive planning. The primary laboratory facility available to the Navy is the Defence and Civil Institute of Environmental Medicine (DCIEM).

Operations. A potentially controversial option is to conduct studies and evaluations during operations. The advantages of this approach is that much, if not all, of data gathering can be

built into the OR system software (invisible to the operators) and that the data collected is the most operationally relevant possible. For this approach to work, it must be determined that data collection does not interfere with the performance of the OR team.

5.2.3 Methods

The previous section reviewed simulation testbeds for examining the performance of humans. This section reviews methods for assessing the performance of DSSs to determine whether they are effective decision making aids.

5.2.3.1 Subjective Methods

A common way to assess the effectiveness of DSSs is by measuring the opinions of the users of a system. For example, user satisfaction is often used as a Measure of Effectiveness (MOE) for DSSs (Cyrus, 1991). The assumption is that the user's satisfaction reflects the usefulness of the DSS.

Subjective measurement does not necessarily mean unsystematic measurement. Adelman (1992, pp. 55-64), for example, describes the Multi-Attribute Utility Theory (MAUT), which can be used to assess the overall utility of a system by having users systematically evaluate attributes of the system. The first step is to decompose the DSS into its component functional features and organize these into a hierarchy of global and specific categories. Users rate the value of each attribute individually. The researcher then applies one of a number of mathematical utility functions to integrate the ratings across all attributes to determine the utility value of the DSS as a whole.

A potential drawback of subjective methods is that it may not be possible to aggregate measures across users. Different users may have different internal scales by which they derive their ratings and there may be no linear transformation that can equate ratings across users.

One potential risk of subjective measures is that success or failure can create a *halo effect* in participants' or raters' minds. Thus, accomplishing a mission can make it appear that participants performed very well, whereas failure can make it appear that participants performed poorly. Both cases can lead to biases in subjective ratings. Of course, success and failure are determined by more than just participants' performance and so ratings should not be based on the outcomes of a test, only the decisions and actions taken.

A number of problems of subjective measurement techniques were noted by Crumley (1988, cited in Webb et al., 1993):

1. Limited number of SMEs available to do the assessment (results may not be generalized with any confidence).
2. SMEs lack specific training to ensure that they score exercises to a consistent criterion.
3. SMEs provide subjective ratings of process and unit performance measures as well as outcome measures (research shows that these sets of judgments are likely to be dependent).
4. Subjective ratings are only provided after the exercise has finished, thereby creating the possibility of a distorted representation of what was

originally perceived based on selective memory for events or features of events.
5. Insufficient exercise scenarios to enable reliability to be appropriately assessed.
6. Controllers of exercises and players provide the ratings, biasing outcome measures.

These problems combine to raise questions regarding the validity, reliability and generality of the results obtained in many studies. Several of these concerns represent pitfalls to be avoided in using subjective ratings for any type of evaluation - not just assessment of battle outcomes.

5.2.3.2 Empirical Methods

A second approach is to assess the performance of users with the DSS to determine whether the DSS has enhanced performance relative to some baseline. This can be done using two broad classes of methods: experiments and quasi-experiments (field studies and case studies) (Adelman, 1992, Ch. 5).

Experiments. Experiments generally follow a *factorial design* in which one or more factors is varied as *independent variables* and other factors are measured as *dependent variables* (see Adelman, 1992; McCall, 1980, pp. 2-3; Ray & Ravizza, 1985). When evaluating a DSS, the independent variables focus on features of the DSS and their availability to the user. The most important independent variable is typically whether participants use the DSS or some existing system. Other factors, however, can be varied, such as training and experience with the DSS. The dependent variables consist of MOEs for assessing how well participants perform tasks. MOEs will be discussed further below (Section 5.3) but they consist of observable aspects of participants' behaviour. For example, common measures include error rate, response time, and number of tasks completed. To be useful, MOEs must directly measure behaviours central to the task and the goals of the DSS.

In addition to the dependent and independent variables, Adelman (1992) argues that care must be taken with four other components of experimental design:

- Participants should be representative of the user population.
- The experimental task should be representative of what users will do in the real world.
- The experimental procedures must not constrain or bias participants' performance with or without the DSS (e.g., suggest a particular strategy, limit options).
- Appropriate statistical analyses should be used to identify significant findings.

Adelman's prescriptions conform to the widely held notion that the experimental task and procedures must capture a realistic working environment. However, just as with cellular automata simulations, there may be value in using more abstract experimental conditions. The goal of an experiment is not to predict future behaviour but to clarify theoretical issues. It is the theory that will allow predictions of behaviour in the field. Thus, any technique that can determine the validity of a theory, whether it reflects realistic conditions or not, is useful in advancing our understanding of performance.

Researchers should be more open to abstract experimental tasks that can address specific theoretical issues provided the links between experiment and theory and between theory and the operational environment are made explicit.

The main features to consider when evaluating experimental design are the validity and reliability of the experiment (Adelman, 1992).

Field studies, Case studies, and Quasi-Experiments. Often, it is not possible to conduct laboratory-based *experiments*, due to constraints of time, money, or the availability of participants. Another common way to examine performance is by conducting structured observations in the field (Adelman, 1992, Ch 5).

A *quasi-experiment* attempts to impose the same level of control as a laboratory experiment in the field. In structure, it is the same as an experiment, with the researcher defining independent and dependent variables, experimental controls, and selecting participants. The quasi-experiment differs from an experiment primarily in the level of control that the researcher has in these areas. Practical constraints will limit control over the selection of participants, randomization, and the control of extraneous variables.

A *field study* is more observational than an experiment. The major problem with field studies is that they violate all types of validity to some extent (Webb et al., 1993). Typically, the researcher is not able to manipulate independent variables but must deal with the conditions that exist in the field (Xiao & Milgram, 1998). Thus, for example, a researcher in the field could not manipulate the kinds of threats encountered by an OR team, the level of risk, the amount of stress of the team, or so on. The researcher may be able to administer dependent variables but may have to employ primarily observational measures if participants cannot, or will not, perform measurement tasks. If this is the case, the researcher is dependent on whatever behaviours are exhibited by participants (Xiao & Milgram, 1998; see also Ray & Ravizza, 1980, pp. 268-270). The conditions in the field can often limit the kinds of observations that can be made.

A *case study* is essentially the same thing as a field study. The term, however, is often reserved for in-depth, observational studies of a single participant. In this case, the researcher has little control but may be able to employ more refined measures tailored to the participant.

5.2.3.3 Validity and Reliability

Key issues determining the value of any evaluation technique are its validity and reliability.

Internal validity. A primary consideration is that the results of the experiment reflect the impact of the independent variables and not extraneous factors (McCall, 1980, pp. 3-4). Internal validity is defined as "establishing the causal relationship between the variables of interest," i.e., the individual and dependent variables (Webb et al., 1993). The presence of extraneous variables that can affect the performance of participants makes it impossible to interpret the effects of the independent variables and make any conclusions on the basis of the experiment (see McCall, 1980, pp. 10-12). In particular, characteristics of the participants (age, experience, maturation, etc.) and of the tasks (workload, instrumentation, etc.) are likely to affect the results of experiments.

Internal validity can be maintained in two ways. First, the experimenter can control factors that might affect performance and specify rival hypotheses. The goal here is to predict all possible effects and incorporate all possible factors in the design of the experiment. The problem is that, realistically, it is impossible to predict or control all possible effects. Consequently, the second means to maintaining internal validity is to randomize extraneous variables across experimental conditions (see Hays, 1988, pp. 52-53). The goal here is to allow chance to roughly equate the impact of extraneous factors across levels of experimental variables so that comparison of experimental variables will not be systematically affected (Ray & Ravizza, 1985, pp. 162-163).

In addition to the risk that extraneous factors may preclude one unambiguously assessing the results of an experiment, there are other threats to internal validity (see Webb et al., 1993):

- **Selection:** Users allotted to different experimental conditions are different as a result of some selection process
- **Maturation:** Users gain some experience or change in some manner during the course of the study
- **Instrumentation:** Changes in the way dependent variables are measured during the study
- **Statistical regression:** Selecting subjects on the basis of extreme pre-test scores using scales of less than perfect reliability
- **History:** Intrusion of an influencing event during the conduct of the testing
- **Inter-subject communication:** If members of the experimental and control groups are allowed to talk to one another during the course of the experiment, they may contaminate any treatment effects, depending upon the information they convey
- **Equalization of treatment:** If research administrators become reluctant to tolerate the differential advantages or disadvantages conveyed upon the experimental group, they may intervene or undermine the procedures which distinguish the two groups
- **Compensatory behaviour by control subjects:** If subjects in the control group perceive that they are receiving a less favorable treatment, they may deliberately control their test performance to reduce or reverse the expected differences, they expect the study to show. The reverse may also occur, when resentment of being placed into what may be regarded as an unfavorable test condition may cause subjects to deliberately perform below their capabilities.

Construct validity. In addition to the internal validity of the experiment, researchers must be concerned that the experiment assess the behaviours it was designed to. Construct validity refers to the development of sound operational measures for the hypotheses being tested (Webb et al., 1993). For example, in evaluating operator performance with and without a new system, the experimenter must be sure that the new system is not systematically associated with any other factor that could affect the measurement device, such as innovative training techniques, enhanced communications, or so on.

External Validity. An experiment is an artificial situation that is meant to inform us concerning theory. The value of the experiment depends, then, on its ability to address general issues. The researcher must ensure that the experiment is consistent with the organizational, procedural, and psychological assumptions underlying the theory (Ray & Ravizza, 1985). If an experiment violates these assumptions then its results will not tell us anything about the theory and will not generalize to situations beyond the experiment.

The external validity is the extent to which results of an experiment generalize to other populations (actual crews) and settings (operational settings) (Webb et al., 1993). The principle factors affecting this is the fidelity of the test conditions to real world conditions; i.e., the extent to which experimental conditions capture the critical aspects of the operational environment and tasks. Assessing external validity requires a good analysis of the operational tasks (see Webb et al., 1993, for a more detailed discussion of this issue).

Statistical Conclusion Validity. Finally, appropriate statistical methods must be used in analyzing data obtained in experiments. Statistical validity is a necessary condition for generalizing experimental results. It does not, however, guarantee meaningful results if the internal, construct, or external validity of the experiment has been compromised.

Reliability. Reliability refers to the replicability of experimental results (see Ray & Ravizza, 1985, pp. 122-124). If experimental results reflect true underlying characteristics of behaviour then the same results should be obtained every time the experiment is repeated (without changes in the experiment). The results of any experiment, however, can be affected by numerous factors, such as experimenter bias, that might affect the outcome. There is, in principle, no way to be sure that extraneous factors have not affected the results of an experiment. Thus, it is important to obtain independent replication of experimental findings to assess their reliability. Unfortunately, experiments are almost never replicated. Instead, researchers typically employ similar methods to extend the questions addressed by research while examining previous issues.

Some critical factors that affect reliability are (see Webb et al., 1993):

1. The precision with which procedures are defined prior to testing: a poorly considered experimental script will allow variations to creep in each time the study is conducted. For example, if explicit instruction concerning how subjects are to perform a task are not given, then high variability in the way subjects approach the task may occur. A frequent situation is the failure to specify whether subjects should place the emphasis on accuracy or than speed when performing the task.
2. The extent to which independent variables (i.e. contrasting experimental conditions) are tightly controlled.
3. The consistency in application of the script by testers
4. The extent to which intervening variables, which could influence test outcome, are anticipated and controlled.
5. The collection of a sufficient number of data points to ensure that the performance of interest is sampled with adequate precision.

6. Consideration and control of temporal order effects which may influence performance, for example, increasing familiarity with the task, or the converse increasing boredom.
7. Measurement precision of the dependent variable.

5.2.4 Level of Fidelity

One issue in developing an evaluation method is the level of fidelity of the system or experimental situation to operational condition (Matthews et al., 1997). It is not always necessary for the test system to look and act like the operational system in all respects. Validity and reliability depend on matching the test and operational systems only on dimensions relevant to the evaluation issues (Matthews et al., 1997). Researchers should carefully consider what elements of the operational system and environment need to be replicated in a study. In general, there are five levels of task complexity for matching test and operational systems (see Matthews et al., 1997):

1. **Part Tasks** comprise the simplest system functions, independent of any scenario. Examples include opening applications, finding files, and selecting buttons. These tasks are normally completed without any direct interaction between individual participants.
2. **Whole Tasks** comprise sub-task components put together in a constrained sequence, in order to complete a specific function. Examples include composing a message and tracking a target. These tasks are normally performed in predefined scenarios without any direct interaction between individual participants.
3. **Task Strings** involve sequential tasks performed to meet a higher-level system goal. An example is calculating an intercept plot. These tasks will be completed in response to predefined scenarios but without any direct interaction between individual participants.
4. **Multiple Tasks** involve extended sequences of tasks and task strings, sometimes performed concurrently, usually over an extended time period. Examples include supervising a target intercept, and planning an operation. These tasks are normally completed in the context of predefined scenarios, but without any direct interaction between individual participants.
5. **Team Tasks** involve sequential and concurrent tasks for individuals performing different roles but working together. An example is mission planning. These tasks are performed in predefined scenarios with direct interaction between individual participants, the overall system under evaluation and other control functions, which may be necessary to fully simulate the operational context.

5.2.5 Scenarios

Studies conducted using simulation and experimental techniques generally rely on the use of scenarios. These are stories or scripts that illustrate some tactical situation or problem. They identify the setting, entities involved, and actions taken by friendly, neutral, and enemy forces. A well-designed scenario will provide opportunities to evaluate the impact of critical issues and factors on human performance (Matthews et al., 1997).

For the naval tactical domain, most scenarios deal with certain kinds of situations (see Hutchins & Kowalski, 1993), including those that involve operation in shallow and confined waters with numerous neutral and enemy aircraft and seacraft in the vicinity. In this respect, the scenarios used in most research are consistent with the current and future naval environment.

The purpose of a scenario is to allow researchers to assess the effectiveness of a human crew, a DSS, or both. Thus, scenarios must present opportunities to measure behaviours at three levels - overall mission effectiveness, team effectiveness, and individual effectiveness (Elliott et al., 1996).

Typically, scenarios should be designed to permit assessment in two broad areas, planning and response. Thus, scenarios will contain a phase during which participants engage in mission planning (Hakala, 1995). This is a deliberate but time-sensitive activity in which the team determines the deployment of resources, interprets ROEs, and pre-plans for certain expected events. Response involves reacting to events, either anticipated or not, that occur in the scenario. In many cases, these events will have been covered by some form of pre-planning. In other cases, the events may have been completely unforeseen. In either event, the team must resolve a crisis situation. A common event used in naval scenarios is *threat assessment*, in which the team is confronted with one or more potentially hostile air or sea units and must assess the identity and intent of the units and determine what actions to take (Hutchins, Morrison & Kelly, 1996; Lipshitz & Shaul, 1997). Although mission planning is an important phase of all operations, not all scenarios deal with it. More commonly, scenarios focus heavily on threat assessment and response.

Scenarios are generally developed in cooperation with operations and HF SMEs (Matthews et al., 1997). This ensures that the scenarios can address both operational and human performance issues. Scenarios development begins with the identification of the mission functions that will be the focus of the evaluation, followed by the identification of human capabilities necessary to accomplish those functions (Matthews et al., 1997). Thus, scenarios rely on task analyses to establish the broad outline of events and human actions.

Several researchers (e.g., Hill, 1996; Mathieson & Miller, 1996) have argued for the need for standardized methods for generating scenarios. Generally, researchers rely on SMEs to guide the development of plausible scenarios. This practice, however, places emphasis on realism rather than experimentally-relevant design. In addition, it emphasizes past experience over current doctrine.

Mathieson and Miller (1996) advocate a study-specific process to characterize relevant scenarios. The goal is to identify the characteristics of scenarios that will likely have a significant impact on the results of the experiment. Analyzing potential scenarios yields a small set of critical features that will drive the answers to the main questions of the study. Thus, any scenario used in the study should contain those features to explore the issues under consideration.

From these features, one creates a *characterization matrix* containing each characteristic as a dimension. The combinations of specific values along all the dimensions define the particular cases or scenarios that can be derived. Mathieson and Miller (1996) developed a set of heuristic guidelines for developing the characterization matrix, which are listed in Table 5.3.

Overall, they argue that a researcher should concentrate on finding the fewest possible characteristics, each of which addresses some critical aspect of the experimental questions.

Rules for Developing a Characterization Matrix	
1.	Characteristics should have direct, relevant effects
2.	Characteristics should be as independent as possible
3.	Characteristics should have few significant values
4.	Many simple characteristics are better than few complex ones
5.	Make maximum use of any study constraints available
6.	Be pragmatic (compromise between simplicity and abstractness and between completeness and usefulness)

Table 5.3 – Rules for Developing a Characterization Matrix (from Mathieson & Miller, 1996)

With the characterization matrix, a researcher can identify existing scenarios that meet the criteria for the experiment and select the ones that best match the questions of the study. If no existing scenario has all the necessary characteristics, the matrix specifies the critical aspects that a new scenario must contain and guides the scenario-writing process. The matrix is also useful for prioritizing scenarios when time or funding are limited.

Hill (1996) has developed the Scenario-based Elicitation Toolset to assist researchers in this development process and it could be used in conjunction with Mathieson and Miller's (1996) characterization matrix approach. The Toolset consists of animatable storyboard software, the main function of which is to structure dialog between SMEs and developers and capture the results of this dialog. There are three main components linked together in an hypertext scenario book.

Scenario Animator. The first component is the scenario animator that is used to graphically represent the region or environment in which the scenario is set and specify the deployment of forces and other important features. Menu options provide means to specify terrain and other physical features and deploy forces for the initial set-up of the scenario.

Communications Animator. The communications animator allows the researcher to set up, animate, and discuss potential communications situations. Using computer technology, the researcher can link communications to animated events. This component contains options for creating communications networks between units within the scenario. These options serve as a focus for discussing communications issues and determining which are most relevant to the study for which the scenario is being generated.

Data Information Organizer. Finally, the Toolset contains a data organizer that provides templates in which data can be entered when the scenario is conducted. Each template records aspects of the scenario elicitation and provides the means by which entities in one template are linked to entities in another. The researcher makes decisions about what to record, how to record it (automated or manual), and when to record it. Overall, the Toolset serves as a computerized guide that researchers and SMEs can use together to create scenarios.

5.2.6 Recommendations

- Conduct an extensive survey of existing laboratory and simulation facilities to identify ones directly applicable to CPF.
- Explore non-traditional methods that allow effective and low cost investigation of theoretically relevant issues.
- Explore cooperative programs with other nations to expand the range of facilities available and learn new methods and techniques.

5.3 Measures of Effectiveness for C2 and Decision Support Systems

5.3.1 Definitional Issues

A crucial aspect of any C2 or DSS design methodology is measuring the effectiveness of the system, where the system is any DSS, equipment, or integrated set of equipment that serves a particular function. Thus, the definition of a system depends on the missions, tasks, functions, and issues under consideration. Measurement of effectiveness, however, raises a number of questions (see Cyrus, 1991). First, what is effectiveness? It is not immediately clear how this concept should be defined or even whether there is a generally applicable definition. Second, once effectiveness is defined, how can it be assessed? Both theoretical and practical considerations should affect the measurement process. Finally, what relationships exist between measurement techniques and other factors affecting performance of a system, such as user characteristics.

The objective of measuring effectiveness is to give an indication of the usefulness of the system to the user, his or her team, their ship, the task group and so on to the national government. Considerable effort will have been devoted to developing and deploying the system and so it must serve needs at all levels of the chain of command. We can identify some specific objectives of measurements in terms of the what they tell us about how well the system meets these needs. Miller, Blanks, & Le (1997) list several objectives of system assessments:

- Determine the utility of the system to the warfighter.
- Identify deficiencies in performing missions.
- Assess C4I system performance in operations.
- Determine training and expertise required to operate the system.
- Collect user feedback.
- Determine the system's compliance to standards.
- Document interoperability (ability to work with other systems).

These assessments must all be made in the contexts of military missions and national policy. Miller et al.'s (1997) analysis points out that an important concern of measurement is to determine whether the system contributes to better overall performance of mission goals by the C2 team. To this extent, measurement is outcome focused. Outcomes, however, are determined by interactions of many factors and so there is a need to isolate the system's contribution in some way. Thus, measurement will also address the operation of the system itself and user's opinions. In this way, measurement is performance oriented. Mason (1995) indicates five broad performance areas that should be assessed:

- System facilitates the user's influence of events.
- System is adaptable to the situation.
- System supports information requirements.
- System is deployable, mobile, flexible, interoperable, and integratable with other systems.
- System supports combat picture effectiveness.

Mason's (1995) performance categories indicate that a system's functionality is largely tied to the human operator's functionality.

A distinction is made between C2 processes and C2 equipment when evaluating systems (Robson, 1997). Typically, MOEs address system processes rather than equipment (computers and other hardware). The role of the system is to support the task at hand. Although the effective operation of hardware is necessary, performance will ultimately be determined by the process components of the system (Robson, 1997). Furthermore, measures developed for processes are more likely to remain current than measures of equipment functioning.

5.3.1.1 Types of Measures

Given the multiple foci of system assessment, researchers have developed multiple kinds of measurements, as summarized by Mason (1995).

Measures of Effectiveness (MOEs). These are assessments of how the system performs functions in the operational environment. Examples include probability of detecting a target, accuracy of detection, and reaction time to specified events. These measures can be applied to actions performed by the system or to the performance of human operators using the system. MOEs are "scenario dependent." That is, they refer to the specific tasks or situations in which measurement occurs.

Measures of Force Effectiveness (MOFEs). These are measures of how a force or operational group performs its mission. This kind of measure assesses the system's contribution to the outcome of the mission. Examples include the success or failure of the force and the degree to which the system met certain mission-defined requirements. MOFEs are also scenario dependent.

Measures of Performance (MOPs). These are measures of properties or characteristics of the system that relate to how well it will meet operational requirements. Examples include throughput, baud rate, and frequency range. MOPs are scenario independent because they assess inherent properties of the system that will exist in all cases. MOPs are related to MOEs but assess characteristics that contribute to effectiveness rather than the level of effectiveness itself.

Physical Parameters. These are measures of the physical characteristics of the system, such as size, weight, capacity, and so on. Physical parameters can indirectly affect how well the system performs its functions. These measures are scenario independent.

5.3.1.2 Measurement Techniques

Because MOEs are scenario dependent, they must be developed with respect to the specific tasks and constraints of the scenario. This means that the operational environment must be specified completely in advance (Mason, 1995). Many useful MOEs have been developed in previous research but they must be adapted for the particular situation under consideration.

MOEs make use of numerous techniques. This is due, in large part, to the many aspects of system performance that researchers wish to assess. Broadly speaking, MOEs employ observational, subjective, and objective techniques (Serfaty et al., 1997; Xiao & Milgram, 1998).

Observational techniques involve monitoring the system's and/or operator's behaviour, typically under operational conditions. Performance-related aspects of behaviour, such as errors, reaction times, and decisions, are recorded and used to determine the effectiveness of the system.

Subjective techniques involve the structured assessment of human operators' opinions, emotional responses, and knowledge states. Subjective techniques focus on determining whether the operator believes the system has performed well or not. These techniques can be especially valuable in cases where the contribution of the system to overall performance is difficult to distinguish from that of the operator or team. In this case, the operator can give some indication of the extent to which the system facilitated performance. Subjective measures can target very specific aspects of performance. This is important because, often, users' overall evaluations are not particularly diagnostic. In addition to assessing the views of users, it is common to employ SMEs to assess systems. Combining observational and subjective techniques allows experts to monitor the effectiveness of the system. Subjective measures can be collected using a variety of procedures, such as questionnaires, interviews, and preference ratings (Miller et al., 1997).

Objective techniques involve the quantitative measurement of discrete, observable behaviours of the system and/or operator, often in laboratory or simulation settings. These measures depend on researchers correctly identifying behaviours that reflect task-relevant processes. In some cases, such as error rate and reaction times, the connection to performance is clear. In other cases, particular theoretical issues will determine the behaviours to be observed (e.g., number of messages sent).

5.3.2 Classification of MOEs

MOEs can also be classified according to the aspect of the system that they measure. Miller et al. (1997) distinguish three classes of MOEs.

Usefulness MOEs. Also referred to as utility MOEs, these assess whether the system provides the required functions to the operator and can perform mission tasks.

Usability MOEs. These assess the ease with which the operator can accomplish tasks using the system.

Performance MOEs. These assess how well the system works in the operational environment, including its efficiency and availability.

Examples of each class are provided in Table 5.4. The remainder of this section describes several specific kinds of MOEs that are illustrative of the subjective, observational, and objective classification as well as the usefulness, usability, and performance classification.

Class	MOE	Description
Usefulness	Value Added	Ability to increase fulfillment of the mission or function
	Completeness	Provide necessary information needed in a timely manner
	Accuracy	Provide error-free data
	Interaction	Permit workstations to collaborate in mission functions
Usability	Human factors	Provide the operator with easy to use and interpret interface
	Consistency	Present like information in the same manner
	Interoperability	Facilitate direct electronic exchange with other systems
	Accessibility	Locate and retrieve information
Performance	Compliance	Compliance with military and commercial standards
	Availability	Ability to be continuously operational
	Bandwidth	Work within limited communications resources
	Year 2000 operational capability	Operate without interruption beyond the year 2000

Table 5.4 – Examples of MOEs (from Miller et al., 1997)

5.3.2.1 User Satisfaction

User satisfaction is a subjective measure related primarily to the usability of the system. It is defined as the user's perception of the quality of the system and the extent to which it meets the user's needs. More precisely, it can be defined as the weighted sum of the user's reactions to features of the system (Bailey & Peterson, 1983, cited in Cyrus, 1991). Measures of user satisfaction depend on identifying the critical features that the operator uses and reacts to. Bailey and Peterson (1983, cited in Cyrus, 1991) compiled 39 general factors for use in a rating scale. This measure, however, is probably most effective if an analysis technique (e.g., factor analysis) is applied to the specific system.

One component of human performance is acceptance of the system. Ives and Olson (1984, cited in Cyrus, 1991) note that user acceptance in and of itself can enhance performance with the system. User acceptance helps the operator develop realistic expectations of what the system can do and fosters commitment by the operator to

learn how to use the system effectively. For this reason, some measures of system effectiveness will focus on the user's subjective evaluation of the system.

Although subjective, the user's satisfaction can affect how the operator uses the system and predict performance (Cyrus,1991). The measure seems to be based on the view that a system that is unacceptable to the user is an unsuccessful system.

5.3.2.2 System Usage

System usage is an objective measure of the usability and usefulness of the system. In this case, the extent to which operators voluntarily use the system indicates how easy the system is to use and how useful it is to the operator (Cyrus, 1991). To operationalize this measure, the researcher should distinguish the way the system is used, whether interactively or off-line (i.e., through an intermediary or through limited output). How the operator interacts with the system will affect opportunities to demonstrate system usage. Also, the researcher must decide on the precise aspect of usage that is relevant to evaluation of the system. For example, usage can be measured by the time spent using the system, by the effort expended while interacting with the system, or the number of outputs obtained from the system.

It may not always be possible to use system usage as a measure for naval C2 system because personnel have limited choice concerning how to perform tasks. Typically, there will be only one kind of system available to perform any given task and operators cannot choose to use a different system in its place. Further, crews are trained to perform tasks in certain ways, with certain equipment. Thus, they will use the system even if it is flawed.

5.3.2.3 Performance

Performance is an objective measure related to the usefulness of the system that can be assessed with respect to the system or user (Cyrus, 1991). It is the measure of the degree to which the task was done properly, regardless of how the task was accomplished (Webb et al., 1993). System performance measures assess performance in terms of resource utilization, efficiency, and cost. User performance measures assess the effectiveness of the user's work with the system in terms of throughput, response times, and reliability.

Use of performance measures require the specification of those aspects of performance most relevant to evaluating the system's usefulness. Leroy and Hofmann (1996) identify attributes related to the quality of information processing and the timeliness of information processing. Attributes related to the *quality* of information processing include:

- Selectivity; the system provides the appropriate amount of needed information.
- Accuracy.
- Ease of comprehension; the system interface facilitates understanding and resolution of errors.

Attributes related to the *timeliness* of information processing include:

- Response time; the system reacts within established time criteria.

- Timeliness; the system presents information at the appropriate times when needed.
- Ease of use; the system provides easy access to information.
- Decision response time; the system allows the operator to make decisions with established time criteria.

These lists of attributes focus primarily on system attributes but also user attributes to some extent, demonstrating the inter-relatedness of these issues.

5.3.2.4 Productivity

Productivity is a observational measure of the performance of the system. It is a measure of how well system resources are used to accomplish specific goals (Cyrus, 1991). In general, productivity can be assessed as the ratio of output to input, where output and input are quantified according to some scale. Often, for example, productivity is calculated by quantifying input and output in dollar values of the materials and labour involved. A dollar scale, however, is probably not as useful for the naval tactical domain as for the procurement and management domains. Research in methodologies should consider the issue of the appropriate way to quantify tactical input and output.

An example of a productivity measure can illustrate how it is used to assess system effectiveness. The *value analysis* method determines what the user is willing to pay in order to keep the system under consideration (Keen, 1981, cited in Cyrus, 1991). Table 5.5 outlines the steps involved (Smith, 1983, cited in Cyrus, 1991). The method calls upon the user to define what features of the system are important. Thus, this is a subjective measure, unlike most productivity measures. Nevertheless, the value analysis method indicates the main feature of productivity measures, that they identify the component contribution of the system to the overall outcome of the task or mission.

Value Analysis Steps	
1.	Define benefits to be obtained if the prototype DSS is developed
2.	Determine the amount of money users are willing to pay to obtain those benefits
3.	Determine whether the prototype can be implemented within the cost threshold established by the user
4.	Design the prototype
5.	Measure the cost and use of the prototype
6.	Review and extend the benefits list if necessary
7.	Define computer hardware and software requirements
8.	Determine the cost of expanding the system
9.	Design a second version of the system
10.	Measure the use and costs and determine the new cost threshold for a possible third stage of development
11.	Continue steps 5-10 until users and designers are satisfied

Table 5.5 – Value Analysis Procedure
(from Smith, 1983; cited in Cyrus, 1991)

5.3.3 Factors Affecting Measures

Measures of user satisfaction, system usage, and performance can all be affected by environmental factors. Webb et al. (1993) identified three classes of such factors:

- Physical Environment: factors such as lighting, workspace layout, temperature, etc.
- Training and Support: level and kind of instruction, training materials, aids, and so on.
- Organization: factors pertaining to the relationship of the individual to the organization and command structure.

5.3.4 Measurement Systems

A number of measurement systems exist for C2 and DSSs. The systems are designed to identify important features of the system and facilitate the measurement of effectiveness. These can be applied, with some modification, to the HALIFAX class.

5.3.4.1 Modular Command and Control Evaluation System (MCES)

MCES is a managerial system designed to guide specification of research problems and guide the analysis of C2 systems (Mason, 1995). It is a methodology that can be used to implement a design process (Section 5.1.2). It consists of seven modules dealing with separate stages in the evaluation process:

- **Problem formulation**: This stage identifies the objective and the precise problem being addressed as well as clarifies the system's operational and deployment concepts, environmental factors, scenarios, etc.
- **System bounding**: This stage uses input from the first module to bound the system in terms of physical entities, structures, and processes, as well as categorize the system elements.
- **Process definition**: This stage builds a dynamic framework that identifies relevant C2 processes, each broken into set of functions, and maps the functions onto a C2 process loop consisting of an environmental stimulus and processes of sensing, assessing, generating, selecting, planning, and directing.
- **Integration of system elements and functions**: This stage integrates system elements defined in the second module and process functions from the third module into a conceptual model of the system.
- **Specification of measures**: This stage identifies MOPs, MOEs, and MOFEs to address problems of interest.
- **Data generation**: This stage gathers data for measures by simulation, exercises, experiments, expert opinion, and survey.
- **Aggregation of measures**: This stage performs analysis and interpretation of observed values of measures.

5.3.4.2 Headquarters Effectiveness Analysis Tool (HEAT)

HEAT was designed to assess overall performance of headquarters but could be modified to deal with the effectiveness of a C2 team (Mason, 1995). Originally, it was developed to determine the optimal size, staffing levels , and distribution of command

posts. To accomplish this goal, it was necessary to define mission goals and functions in detail (Webb et al., 1993). HEAT uses a decision making model to aid the design of MOEs and develop a plan for collecting measures. The model considers six steps in the decision making process (Mason, 1995):

- Monitoring (obtain data).
- Estimation (place values on monitoring).
- SA.
- Opinion generation (evaluate the feasibility of COAs).
- Opinion selection.
- Plan generation/direction.

Plan generation is the most important step because HEAT evaluates performance primarily in terms of the quality of the plans generated (Webb et al., 1993). In general, effectiveness is assessed in terms of the magnitude of changes or revisions required for plans over their intended life span; i.e. by the effectiveness of the planning activities that generated them.

MOEs can be developed for each of these steps to provide a complete assessment of the effectiveness of decision making.[6] The MOEs developed in HEAT are defined under four types:

- Characteristic measures.
- Coordination measures.
- Queuing measures (e.g., communication).
- Quality measures (Performance).

HEAT relies on a pre-defined set of MOEs, which may limit its generalizability to new situations.

5.3.4.3 Army Command and Control Evaluation System (ACCES)

ACCES is a comprehensive methodology for evaluating C2 functions. It was developed in the context of army C2 but can be adapted to naval tactical C2.

ACCES is a framework for analyzing events in the context of exercise scenarios (Webb et al., 1993). It focuses on thee tasks people perform and their goals and information needs. So far, ACCES has been applied to six major areas (Keener et al., 1990, cited in Webb et al., 1993):

- To build C2 databases.
- To provide exercise feedback.
- To provide training evaluation and development.
- To evaluate equipment use.
- To conduct C2 experiments.
- To support simulation development.

[6] The assumed model is a hybrid model, combining analytic and intuitive processes, and may not be completely applicable in all situations.

Webb et al. (1993) point out that ACCES can also be used to:

- Evaluate system design alternatives.
- Evaluate alternative approaches to organization structure and function and task allocation.

The ACCES methodology divides the activities of C2 staff into nine major areas (Webb et al., 1993):

1. Collecting information through monitoring the environment and receiving reports.
2. Synthesizing and analyzing.
3. Developing recommended action alternatives.
4. Reviewing recommended action alternatives.
5. Planning implementation.
6. Reporting.
7. Coordinating.
8. Seeking information.
9. Disseminating information.

From these activities, ACCES identifies certain observable elements that can be used to characterize C2 performance (Webb et al., 1993):

1. Information about the environment.
2. An initial understanding of the environment.
3. An overall estimate of the alternate actions required, their anticipated results, and consequent recommendations.
4. Specific decisions of the commander and others in the command chain.
5. Reports that inform others.
6. Command guidance.
7. Plans and directives.

The observable elements are used to determine the effectiveness of the C2 system, where effectiveness is defined as the development of stable plans, issuing directives accurately, executing directives accurately, and responding to changes in situation quickly (see Webb et al., 1993, Annex B for further description of the ACCES process).

5.3.4.4 Functional Decompositions

Another approach is to allow SMEs to define the critical features of the system and task upon which MOEs are based (Kemple, Stephens & Crolotte, 1980, cited in Mason, 1995). The benefit of this process is that experts will presumably have the best knowledge of how the system is, or will be, used in the field. Value analysis, discussed previously (Section 5.3.2.4), is one such technique.

In a functional decomposition, SMEs divide the domain of the system into three categories: mission, organization, and resources (Girard, 1989, cited in Mason, 1995). For each area, the SMEs generate Mission Success Criteria (MSC) that indicate the performance components of the system necessary to achieve the mission, and Required Capabilities (RCs) criteria that are used to generate functions. Specific MOEs are then generated around MSC and RCs to create a set of measures addressing critical functional components of the system.

5.3.4.5 Multi-Attribute Utility Theory (MAUT)

MAUT is based, to some extent, on the SHOR framework (Wohl, 1981, cited in Adelman, 1992; Wohl et al., 1988). The steps of SHOR, stimulus, hypotheses, options, and response, involve interactions of a number of entities, including the human decision maker, the organization, and the environment. Thus, the SHOR framework defines three areas of measurement for determining system effectiveness (Webb et al., 1993):

- The interface between the system and user.
- The interface between the user and organization.
- The interface between the organization and environment.

Each interface contributes to the system's effectiveness and requires its own evaluation criteria. The first step of the MAUT methodology is to identify MOEs specific to each interface. To do this, each interface is analyzed into its component processes, functions, and activities (Webb et al., 1993).

The next step is to identify evaluation metrics that assess the effectiveness of each component process. These measures are domain specific, although they generally refer to several important aspects of system processes – time, accuracy, and usability. The MAUT hierarchy of evaluation criteria are listed in Table 5.6.

All the evaluations of component processes are done within the context of the overall hierarchical organization of those processes. In other words, components are linked in a weighted relational tree that indicates which processes contribute to others and to what extent. The final step of MAUT is to aggregate upwards the weighted scores in the hierarchy to generate composite scores for each process and each interface domain, then an overall composite score for the entire system utility (Webb et al., 1993).

5.3.4.6 Multi-Criteria Analysis (MCA)

MCA assesses effectiveness in terms of the system's contribution to overall mission success (Robson, 1997). The approach is highly analytic, employing tools to decompose the system into component functions or options and comparing the contribution of each to the achievement of mission objectives. The aim of MCA is to develop a model of the system that identifies the relative influence of component functions in terms of C2 activities and characteristics.

Overall Utility		
DSS-User-Interface	**User-DSS-Organization**	**Organization-Environment**
Match with personnel	Efficiency factors	Decision accuracy
Training and technical	Time	Match between DSS' technical approach and problem's requirements
Work style, workload, and interest	Task accomplishment	Decision process quality
Operational needs	Data management	Quality of framework for incorporating judgment
DSS' characteristics	Setup requirements	Range of alternatives
General	Perceived reliability under average battle conditions	Range of objectives
Ease of use	Skill availability	Weighing of consequences of alternatives
Understanding	Hardware availability	Assessment of consequences of alternatives
Ease of training	Match with organizational factors	Reexamination of decision making process
Response time	Effect on organizational procedures and structure	Use of information
Specific	Effect on other people's position in the organization	Consideration of implementation and contingency plans
User interface	Political acceptability	Effect on group discussions
Data files	Other people's workload	Effect on decision makers' confidence
Expert judgments	Effect on information flow	
Ability to modify judgments	Side effects	
Automatic calculations	Value in performing other tasks	
Graphs	Value to Strategic Air Command or other services	
Printouts	Training Value	
Text		

Table 5.6 – MAUT Hierarchy of Evaluation Criteria (from Adelman & Donnell, 1986, cited in Webb et al., 1993)

MCA is similar to MAUT. It employs a multi-attribute rating technique to gather data concerning the utility of system features. The evaluation process follows the following steps (Robson, 1997):

1. Identify the functions and objects of the evaluation.
2. Identify the stakeholders of the system.
3. Identify the attributes (functions, options, etc.) of the system and organize into a *Value Tree*.
4. Obtain stakeholder assessments of the relative importance of all attributes (ratings).
5. Assign linear weights based on the assessments.
6. Aggregate value data to obtain an overall system value.
7. Perform a *Sensitivity Analysis* to indicate the robustness of the model.

The assignment of weights to the stakeholder assessments can be done by a "swing method" (Robson, 1997). This involves the comparison between a swing from the least preferred to most preferred value in one attribute to a similar swing in another.

The evaluation of the system depends on a mathematical value function that converts the weights into a value score. The value function represents how stakeholder assessments are affected by variations in values to a particular attribute.

MCA can be applied at various levels of analysis, including mission, system, subsystem, and equipment. The method, however, requires development of a set of metrics at each level. Metrics at the mission level include percentage of enemy forces killed and percentage of own forces killed. Metrics at the system level, which is perhaps the level most relevant to tactical C2 systems, include:

- Tactical picture compilation.
- SA.
- Resource allocation.
- Resource direction.
- Planning.
- HCI.
- Communications.
- Intelligence.

5.3.5 MOEs for Military Operations Other Than War (MOOTW)

Most research has considered MOEs in the context of warfare and traditional military operations. With changes in the role of the Navy, it is becoming increasingly important to assess the effectiveness of C2 and DSSs in the context of non-combat missions (Wall, 1997). These tend to be protracted and complex missions and require the interaction with civilian agencies. Thus, C2 systems will be called upon for use in situations for which they were not designed. In particular, MOOTWs add complexities to the question of what constitutes an effective system (Wall, 1997):

- Long duration of operations requires extended operation of the system.
- Restraint and security require the deployment of force, including security, while minimizing friction and escalation and maximizing cooperation and legitimacy.
- Multi-agent participation requires interaction, often in a support role, with civilian agencies.

- High political content and objectives of operations requires the inclusion of legal, ethical, political, and other non-military issues in the decision making process.
- Absence of simple solutions requires the commander to seek information from many sources, including sources not traditional in the combat role.

Wall (1997) argues that MOEs for evaluating C2 systems in MOOTWs should:

- Reflect criteria for mission success.
- Help commanders assess the readiness and ability of other agencies to take over responsibility.
- Serve as a basis for comparing the relative merits of alternate COAs.
- Be reasonable in number.
- Be sensitive to changes in the measured variable without being unduly affected by extraneous influences.
- Be based on a broad base of knowledge and take advantage of the experience of other agencies.
- Be objective or have clear caveats/conditions.
- Help commanders assess changes in the operational environment (i.e., measure relative to culminating points in the mission, assess progress by milestones, etc.).
- Be appropriate to the concept of the operations.

In addition, Wall (1997) suggests seven categories of MOEs that are related to Humanitarian Aid missions in particular (specific MOEs would have to be developed on a mission-by-mission basis):

- Concept of operation development (participating agencies meet to agree on mission goals, and COA).
- Derivation of task listing (detailed plan of action).
- Development of criteria (decide what data is required to support decision making).
- Development of variables (decide what variables must be measured to assess criteria).
- Development of methodology.
- Execution and continuous reassessment.

5.3.6 Recommendations

- Address definitional issues of measurement early, targeting issues of how the system is to perform, what it is to achieve, and what aspects of the human user's performance that the system is designed to enhance.
- Develop MOEs for MOOTW.

5.4 Analysis Techniques

In addition to experimental and field study techniques, researchers employ a range of analysis techniques designed to describe complex behaviour. This section will review three popular analytic techniques for characterizing human performance and decision making, namely Cognitive Task Analysis (CTA), Cognitive Work Analysis (CWA), and Conceptual Mapping (CM).

5.4.1 Cognitive Task Analysis

CTA is actually a general term for a family of methods designed to analyze the thought processes of people while they do a job or task (Randel et al., 1996). These techniques contrast with task analysis in their focus on cognition rather than behaviour in the task. Task analysis is a procedure to determine (Luczak, 1997):

- The elements comprising a task.
- How elements are logically arranged.
- How elements are sequenced in time.
- How elements or the task itself can be explained or justified.
- How elements relate to other tasks.

Thus, task analysis establishes and characterizes the behavioural steps involved in performing a task. As such, it is a precursor to CTA. It is important in CTA to identify the behavioural components of the task because these will guide and influence cognition. The goal of CTA, however, is to identify the knowledge, skills, decisions, and information processing involved in performing the task. In this respect, CTA builds on task analysis.

CTA is a useful technique when the task is complex, difficult to learn, dynamic, and ill-structured (Gott, 1994, cited in Randel et al., 1994). These criteria apply to the naval C2 domain where much of the task-related activity occurs in the operator's head. For these sorts of tasks, CTA can identify the information needs, strategies, and cognitive skills needed to perform the task. These are crucial to characterize not only training requirements but also decision support needs. In fact, the DSS design methods discussed in Section 5.1 assume the use of some method to identify information needs and cognitive strategies.

Although there are variants of CTA, the general method involves the use of structured interviews based on standardized scenarios. Scenario selection, as we will see, is important for determining the generalizability and applicability of the results to the research goals, especially in the naval context (see also Section 5.2.3). Scenarios should be developed in conjunction with SMEs who can ensure realism in content, pacing, workload, and data/information. During scenario development some idea of the subtasks or steps involved in the overall task is developed and used to structure the interview protocols for participants.

After development, scenarios should also be validated with independent SMEs. Relatively little work has empirically validated CTA techniques and there are no agreed upon criteria or validation techniques (Cooke, 1994, cited in Hoffman et al., 1998). In fact, there are not necessarily many common features on which to compare CTA methods. Thus, validation must proceed on an individual basis.

Once scenarios have been chosen, CTA follows four general steps (Webb et al., 1993; Miller et al., 1992; Klein, 1997):

1. **Initial briefing**: Outline the purpose of the study to participants, explain key terms, and obtain correlational information (level of experience, age, etc.).
2. **Knowledge elicitation**: Using schematic representations of key working environment resources (equipment, etc.), prompt participants to describe the types and sources of information needed at various stages of the task.
3. **Detailed discussion of the scenario**: Step participants through the scenario, probing for specific information needs, decisions, strategies, and task goals, and

record data in terms of the cognitive skills demonstrated (SA, decision making, communication, workload, etc.).

4. **Debriefing**: Review the results of step 3 with participants and confirm and elaborate the data.

There are a number of variants on this general method. Each specific method addresses somewhat different aspects of cognition in somewhat different ways. The following sections review some of these CTA methods.

5.4.1.1 Task Process Model (TPM)

Randel et al. (1996) reviewed a number of CTA methods, including TPM. The goal of TPM is to create a graphic or schematic representation of the task. Thus, it focuses more on the logical structure of the task than other CTA methods. The representation of the task, however, is derived from interviews with experienced performers and summarizes their cognitive approach to the work.

The researcher prepares for knowledge elicitation by:

- Collecting and reviewing relevant documentation (training manuals, requirements, etc.).
- Observing training exercises and/or actual job performance.
- Selecting participants who are SMEs in the domain.

The researcher then conducts interviews to a) develop a model of the task, and b) identify the critical components of participants' decision making process. Interviews are conducted in two stages. The first interview is to develop the preliminary model and the second is to verify the model and ensure the accuracy of the data collected.

The Stage 1 interview is conducted during or immediately after a simulation, training exercise, or description of a critical event. In all case, the scenario should contain all the usual elements of the job as well as relevant special events. Calderwood et al. (1987, cited in Randel et al., 1996) developed a Critical Incident Interview Guide to assist in developing interview techniques to identify critical incidents. The interview focuses on decision points. For each decision point, the interview elicits data concerning the cues used by the participant, errors made, missing data, differences between experts and novices, knowledge and experienced used, decision options, and rules of thumb (see Randel et al., 1996). Three to five participants are required at this stage.

From the results of the Stage 1 interview, the researcher develops preliminary models for each participant by drawing task process flowcharts. The researcher then creates a composite model by combining the individual flowcharts. The composite model is based on the elements mentioned most often and elements that were infrequently mentioned but marked as critical. Thus, the model building process is somewhat subjective and relies on the researcher's interpretation of the data. A valuable modification to this technique might be to involve independent SMEs in the development of the composite model to verify critical elements.

The Stage 2 interview involves an additional 8 to 20 participants, preferably a mix of experts and novices. The general procedure follows that of Stage 1 but new questions are added based on the results of the earlier interview. New task process models are developed as before and a second composite model created by combining individual models. The composite model is validated by independent SMEs (e.g., instructors,

experienced personnel) who indicate the accuracy of the information content, processes, and sequencing.

5.4.1.2 Information Flow Model (IFM)

The IFM method (see Randel et al., 1996) creates a map of the information processing during a team task. The goal is to create a flow chart that depicts who receives information, who sends it, and what transformations occur at each step.

The model is developed by interviewing participants about a scenario. Participants draw diagrams of how information would be communicated in that situation. Each individual in a team task (i.e., an information receiver/transmitter) is represented by a box. Arrows indicate the direction of information transmission. The IFM method works more effectively if the researcher employs a number of representative scenarios.

After interviewing participants, the researcher creates a composite diagram containing all the elements mentioned by participants. Again, this is a somewhat subjective procedure.

The value of the IFM method is that the model indicates both direct and indirect communication. It is very useful for comparing information needs of individuals against communication procedures to determine how well those information needs are met.

5.4.1.3 Misconceptions Analysis

Misconceptions analysis (Randel et al., 1996) documents the rules and heuristics used by people in their work. It addresses this issue by identifying the ways in which people misunderstand their tasks. Thus, this method can be helpful in identifying inefficient or problematic strategies. It is also useful for identifying organizational factors that induce or create the climate for poor work strategies to emerge.

The goal of misconceptions analysis is to identify the rules of thumb used by individuals and the conditions in which these rules apply. Problems can arise when novices do not understand the boundary conditions of heuristics and attempt to misapply them.

Misconceptions analysis has three broad steps. In the first step, the researcher asks participants what advice or rules of thumb they would give to a new worker. The researcher attempts to obtain a list of 12 to 15 such heuristics. Participants rate the usefulness or importance of each rule and the researcher selects the most important. This procedure should involve both expert and novice participants to obtain a range of heuristics.

In the second step, the researcher asks participants to give reasons why each rule is good or bad. The researcher notes the conditions or situations for applying each rule.

In the last step, the researcher points out contradictions between rules generated by the participant and asks for clarification. The researcher attempts to determine whether there is a real contradiction and, if so, how the contradiction is resolved. This step determines the precedent of rules.

Once participants have generated the rules of thumb, the researcher represents the composite knowledge base by:

- Listing the rules that are generally considered useful, including elaborations and exceptions.
- Listing incompatible rules, including explanations and conditions for the use of rules.

The results of misconception analysis are sets of "working practices" that partially represent the way people do their job and their conceptual understanding of the task.

5.4.1.4 Critical Incident Method (CIM)

CIM, also called Critical Decision Method (CDM), was developed by Klein and colleagues (Klein, 1992; Klein, Kaempf, Wolf, Thorsden, & Miller, 1997; Miller et al., 1992) based on the RPD model. Thus, CIM is based on the theoretical premise that performance derives from schematic knowledge that is organized around problem types. The goal of CIM is to describe the nature performers' schemata (Hoffman et al., 1998). To do this, the technique relies on "keys" (probes, queries, etc.) to unlock the performer's implicit knowledge and cognitive processing to produce an explicit record.

This method differs from most CTA techniques in that it does not employ scenarios. Rather, participants generate previous events from memory and these are used as the focus of analysis (Klein, 1992; Miller et al., 1992). The reason for this is the assumption that most expertise develops during critical events rather than routine ones. Thus, CIM uses participants' judgments to identify the most important events rather than relying on scenario generation by other SMEs.

Another distinctive feature of CIM is its focus on identifying Decision Requirements (DRs) (Klein et al., 1997). DRs consist of key decisions that must be made during a task and the way those decisions are made. Thus, DRs are very specific task components required by the task. Identifying DRs is a major part of modeling the cognitive processes of accomplishing the task.

The CIM is conducted by interviewing participants retrospectively about the decisions they made during non-routine events that occurred during actual performance of the task in real-world working conditions (Klein et al., 1997). The goal of the interview is to chart the critical event by time and sequence of decisions and to identify the cues and strategies used by participants at different points in time/sequence. The representation of cognitive strategies can later be reviewed to trace the development of SA, formulation of plans, and the application of expertise and knowledge.

CIM follows eight steps (see Hoffman et al., 1998 for a full description of the steps):

1. Researcher acquires domain knowledge (learn about systems, tasks, etc.).
2. Knowledge elicitation (interview SMEs to obtain critical incidents; use queries and probes to elicit detailed information about their schematic knowledge).
3. Extract incidents (review with SME the interview transcripts to identify event structure, information requirements, and DRs).
4. Identify actions.
5. Identify SA.
6. Identify cues, factors, processes that contribute to SA and performance.
7. Identify critical problems.
8. Code processes and strategies used to create and maintain SA, make decisions, and generate a COA.

The preparation step is needed to identify the goals and constraints of the task. The researcher can then prompt participants for data regarding these during the interviews. Prior to the interviews, the researcher creates a table to record key data elements, specifically, cues, factors, processes, SA, and DRs (Miller et al., 1997). The interviewer should elicit progressively deeper and more detailed data about SMEs' knowledge structure through steps 4 to 7.

Data is represented in several ways. Interviews are audiotaped to preserve raw data and the audiotapes are transcribed (typically requiring 30 to 40 pages) to create a more analyzable format (Klein, 1992). The researcher reviews the transcripts to derive brief accounts that describe the essential aspects of the incident reported by the participant. The data is then converted to a graphic representation, such as a *decision flow diagram*, to describe how the participant thinks about the task (Klein, 1992). The decision flow depicts the DRs and links them to critical patterns of cues and necessary information.

Primary Goal States of AAW Teams
• Determine intent
• Recognition of a problem
• Take actions to avoid escalation
• Take actions toward engaging tracks
• Monitor on-going situation
• Identify track
• Allocate resources
• Prepare self-defence
• Conduct all-out engagement
• Monitor tracks of interest
• Reset resources
• Collect intelligence
• Trouble-shoot
• Determine location of track
• Other (miscellaneous goals)

Table 5.7 - Primary Goal States of the AAW Team (from Miller et al., 1997)

There are numerous examples of this method. Miller et al. (1997) employed CIM to study the performance of Anti-Air Warfare (AAW) teams aboard AEGIS class cruisers (see also Randel, Pugh, & Reed, 1996). From the initial familiarization, the researchers developed a table of problems that defined the goals of the AAW task. These problems included:

- Determine the intent of unknown tracks.
- Deal with tracks that are hostile.
- Tracks become increasingly hostile.
- Tracks have gotten too close.

During interviews, Miller et al. (1997) identified associated SA, DRs, strategies, and COAs that were used to address the problems. The primary goal states of the AAW team were identified (see Table 5.7) and incidents were coded on a timeline to indicate when goal states were active and for how long. The most frequent goals were, first and foremost, identifying tracks, then recognizing a problem, actions toward engagement, and engaging tracks. The researchers inventoried the cues used in each critical incident (see Table 5.8 for cues used in a harassing F-4 incident). The analysis accomplished two things - the identification of decision points and the identification of cues, knowledge, and strategies needed to make the decisions (see Randel, Pugh, & Reed, 1996).

Cue Inventory
• Intelligence reports
• Recent hostilities/activities
• Presence of a new track
• Course intercept and circling
• Range of tracks to ownship
• Point of origin
• Change in range of tracks to ownship
• EW emissions
• Change in course
• Flight profiles

Table 5.8 – Primary Cues Used in a Harassing F-4 Incident (from Miller et al., 1997)

A number of concerns have been raised about the validity of CIM (Hoffman et al., 1998). Wilson and Schooler (1991, cited in Hoffman et al., 1998) have argued that think-aloud protocols such as this can alter how people think and act with respect to a task (but see Ericsson & Simon, 1993, cited in Hoffman et al., 1998). Whether this is truly a problem may depend on the particular participants' verbal and observational skills but care should be taken that queries and probes do not suggest to participants how to approach the task. The questions asked can certainly bias recall of critical events (Loftus, 1980). Although it is not clear that any CTA technique is at substantial risk for this, CIM may systematically distort participants' memory if questions are not carefully developed.

5.4.2 Strengths and Weaknesses of CTA

CTA is widely used and has a number of advantages as a technique to characterize human performance. First, it is "data rich." That is, CTA obtains a great deal of detailed information about how people perform their jobs. Second, the data gathered relates to concrete, realistic events. Thus, the results of CTA are likely to be immediately applicable to Research and Development (R&D) efforts.

One weakness of CTA is that it is a labour and time intensive technique that generates a large volume of data (Kirschenbaum, Gray & Ehret, 1998). As such, researchers can rarely conduct CTA with more than a handful of participants, potentially limiting the generalizability of results. Another weakness is that CTA involves subjective interpretation of the data. Researchers must decompose participants' responses and formulate a general task flow largely by intuition. Thus, there is a risk for bias in recording and interpreting data (Kirschenbaum et al., 1998).

Klein's CIM technique was designed to address one problem of other CTA techniques, namely that of scenario generation (Miller et al., 1992). Scenarios used in CTA will exert a large effect on what results are obtained. To the extent scenarios do not reflect real or important task situations, the results of CTA will be biased. CIM is based on an analysis of events that actually happened and so has somewhat greater validity and generalizability. In addition, CIM can elicit a wider range of contexts and problems than CTA methods that present a set of fixed scenarios (Miller et al., 1992).

A major problem with CIM, however, is that it relies on participants' memory for events that may have occurred months or even years ago (Miller et al., 1992). Given that human memory is highly integrative and combines experiences into schematic representations of past events (e.g., Loftus, 1980), it is likely that participants will not have true memory of specific events. Instead, they will be able to recall generalized, ideal cases of classes of events.

In addition, it is impossible to determine how well incidents recalled by participants represent real world conditions (Miller et al., 1992). Other factors than the criticality of events will affect participants' recall. In fact, people are more likely to remember unusual, distinctive events in detail than normal, common events (Zechmeister & Nyberg, 1982, pp. 270-272). This is useful in identifying critical events but can bias participants to recall situations that do not capture routine but important situations.

Finally, CIM will necessarily emphasize strategies and DRs currently in use (Miller et al., 1992). Because participants will be remembering past events, where they used existing equipment, doctrine, and practices, there is no opportunity to consider new strategies that might be enabled by changes in C2 equipment or doctrine. This limits the usefulness of CIM for research examining new DSSs.

5.4.3 Cognitive Work Analysis

CWA is a new paradigm for analyzing the collaborative control of systems by humans and machines (Vicente, in press). This framework addresses more than just the actions of an operator performing some task with a given piece of equipment; it attempts to describe the entire work domain in which the operator works.

Four layers of aspects or components can describe complex systems:

- Environmental context.
- Organization/Management.
- Workers.
- Technical System.

All layers contribute to the function of the systems and so an analysis paradigm should address all layers. Such an approach is necessary because of the complexity of technical

systems. In particular, a number of factors contribute to the problem of controlling technical systems:

- Large problem spaces.
- Social components.
- Dynamic systems.
- Automation.

Perhaps the most important factor contributing to the complexity of systems, however, is the presence of *disturbances* that affect how well the system functions. Disturbances are unexpected events that affect the operation of the system. They cannot be predicted ahead of time and are part of the environment (physical and social) in which the system operates. Control of the system is threatened when it is not sufficiently flexible to allow a response to the disturbance.

To deal with these multiple levels of complexity, researchers can use work analysis, a technique for examining the environmental, organizational, human, and technical aspects of a work domain. This technique distinguishes constraints on how work can be performed at all levels. Cognitive constraints, for example, originate with the worker and consist of aspects of human cognition that determine what a person is able to do. Environmental constraints are demands that originate in the context in which work is done. Work analysis is based on the ecological approach (Section 2.6.1) and gives primary importance to constraints at the environmental level. CTA, in contrast, gives more weight to cognitive constraints. Thus, work analysis with CTA can provide a richer description of a work domain than either technique alone.

CWA is a new paradigm that extends work analysis and incorporates CTA. The goal of this paradigm is not do describe how people currently perform tasks but to describe the scope of possible work practices that could be used to achieve the goals of the system. Thus, CWA identifies the requirements that need to be fulfilled to support effective work and the constraints that limit the scope of possible practices.

CWA targets five aspects of work:

- Work domain: the system being controlled.
- Control tasks: the goals to be achieved.
- Strategies: the generative mechanisms for achieving goals.
- Social organization: the relationships between individuals in the system.
- Worker competencies: cognitive constraints.

These aspects are ordered by their importance to the functioning of the system. At the highest level, the work domain establishes the principles of the work to be done, including the goals to be achieved. Thus, this level affects all other levels. Each transition between levels represents a step away from ecological constraints and toward cognitive constraints. CWA assumes that cognitive constraints are meaningful only in the context of ecological constraints.

Each level also represents a phase of CWA. The first phase is to show what the controlled system is capable of regardless of the control tasks, social organization, etc. This is a description of the formal structure of the domain and the constraints that, by principle, limit what can be done. The second phase is based in the fundamental constraints of the work domain and identifies additional constraints arising from the kinds of possible control tasks

that can be used to manage the system. Because there are typically a large number of possible control tasks to achieve the same goals, these constraints are described at a high level of abstraction. The third phase is to identify possible strategies that are consistent with domain and control task constraints. The fourth phase is to analyze constraints imposed by necessary human relationships within the context of strategies and control tasks. The final phase is to identify cognitive constraints and how these affect strategies.

CWA approaches numerous design issues. In particular, it can be used to identify constraints that can then be embedded in the design of a system so that operators have enough flexibility to adapt within the remaining space of possibilities. In this way, designers can let workers "finish the design" by managing the system in response to factors that could not be anticipated by designers. Thus, the ecological perspective of CWA suggests more distributed control of systems, with less control being exerted by designers through the choice of a particular set of tasks and strategies.

5.4.4 Conceptual Mapping Techniques

In addition to CTA, conceptual mapping techniques are popular means to analyze and characterize performers' mental representations of their jobs. These techniques do not identify the actual cognitive steps involved in performing tasks. Rather, they provide a more detailed description of the participants' knowledge, including the organization of task-related concepts. This knowledge representation can be used to identify information requirements as well as conceptual relationships that facilitate or impair the use of certain strategies.

5.4.4.1 Card Sorting

Card sorting is often used to identify the relational structure of concepts in participants' minds (see Federico, 1995). The researcher first identifies a set of task-related concepts, typically in consultation with SMEs who indicate (see Klein, 1989, 1993):

- Critical cues.
- Situational expectancies.
- Plausible goals.
- Typical actions.
- Other factors relevant to the particular task.

These concepts are printed on cards and given to participants who are asked to sort the cards into two or more piles by placing similar concepts together. Participants are asked maximize the similarities among concepts in a pile and are free to create as many piles as they like. Overlapping or nested categories, however, are not allowed.

After sorting the cards, participants label the categories. Often, participants will be asked to re-sort the cards to provide a reliability check (Federico, 1995). Also, the researcher interviews participants about the bases they used to sort the cards, indicating, for example, whether participants relied on surface or perceptual features of the concepts or underlying functional similarities.

Finally, participants' sorted categories are used as data in some scaling technique such as Multi-Dimensional Scaling (MDS), Cluster Analysis, or so on (Kruskal & Wish,

1978). The output of these analyses is a graphical representation of the similarity relationships between concepts.

5.4.4.2 Structural Knowledge

A technique related to card sorting is to obtain pairwise similarity ratings for concepts (Randel et al., 1996). Concepts are generated as in Card Sorting. Participants are then presented with every possible pairing of concepts. For each pair, participants give a rating of similarity using some standardized scale. The similarity judgments are then used in MDS or other techniques to depict the conceptual space of participants. Randel et al. (1996) recommend the use of *Pathfinder Net*, a software package that generates a graphic network from similarity data. The output depicts concepts as nodes and relations as paths. The paths indicate which concepts are directly related such that the length of the path indicates the strength of the relation as inferred from the similarity data.

5.4.4.3 Conceptual Graph Analysis (CGA)

Another similar technique is CGA (Gordon & Gill, 1992, cited in Militello, 1998). Like card sorting and structural knowledge techniques, CGA generates a graphical depiction of participants' task-related knowledge. CGA, however, focuses on the task decomposition. The task is broken into its component parts, goals, knowledge, actions, and so on. During interviews, participants are probed to elicit the components of the task and their relationships. From the interviews, the researcher generates a map representation consisting of nodes (concepts) and links (relationships). This map depicts how participants mentally represent what they do in the task, including the flows of information and action.

5.4.5 Recommendations

- Expand the use of ecological approaches (e.g., CWA) that can provide insight into environmental and organizational constraints on how work is performed.
- Create a CTA method specifically for the CPF upgrade from desirable elements of existing methods to accommodate the team environment, complex interactions and communication, and assess KSAs and shared KSAs.
- Study ways to address problems of current CTA methods, notably the subjectiveness of data interpretation and the costs in terms of labour and time.

5.5 Team Research Techniques

Studying teamwork and team performance poses unique problems for researchers. There has been a great deal of research examining work group effectiveness but much of this has been done in the context of private industry (Cannon-Bowers et al., 1992). Although this research has developed team methodologies, much of this work is not directly applicable to military situations. Most industry teams have considerably less hierarchical structure than military teams. Also, there is a greater focus on solving problems through consensus building.

5.5.1 Issues

In designing measures of team performance, there are at least three questions that must be resolved (Baker & Salas, 1997):

- What to measure?
- When to measure?
- How to measure?

There are theoretical, methodological, and practical issues surrounding all of these questions. In answering the first question, the researcher must identify the appropriate behaviours, decisions, KSAs, and so on to be measured, as well as the appropriate units of measurements and whether to target individual members or the whole team as a unit. In addressing the second question, the researcher must consider that teams change and mature over time (Baker & Salas, 1992, 1997) and identify the relevant stages of team development. Finally, when addressing the third question, the researcher must determine how team characteristics can be quantified. This is generally a difficult thing to do and many studies have relied on the use of observational techniques and raters to assign values to measures (Baker & Salas, 1992). This technique can be problematic because the critical skill dimensions must be identified and clearly presented for raters to be consistent.[7]

Unfortunately, skill dimensions for measurement can rarely be defined objectively. Instead, they are generated by theories of teamwork. This can cause the measure to contribute to a "self-fulfilling prophecy" in which the researcher only measures those characteristics a particular theory can predict and ignores other important factors, leading to the mistaken belief that the original theory had been confirmed. Raters themselves may be biased theoretical concepts in their assessment of team concepts (Rouse et al., 1992).

There are other methodological issues related to team performance. First, team performance is partially a function of individual performance and so researchers must be able to distinguish individual and inter-relational factors contributing to team performance (Elliot et al., 1996). In addition, because team members are so highly interdependent, the performance of any individual may be difficult to measure. The accumulation and escalation of error can mask the locus of a problem.

Second, measures of team characteristics and attitudes may not be conceptually simple entities. It may not be possible to measure the team by the sum or average of individual characteristics (Elliot et al., 1996). If this is the case, the researcher must develop measures that capture the entire team. Aside from performance measures (e.g., accuracy, team response time), this can be practically difficult.

A third potential problem is that researchers are often have limited access to participants. This can be especially problematic if researchers are unable to assemble a complete team. Although small group research has been criticized as being artificial, Driskell and Salas (1992) note that this is a problem of all laboratory research. They argue that as long as the study has been properly designed to address theoretical issues, small groups will be sufficient to replicate theoretically relevant aspects of the real-world situation. Nevertheless, small

[7] These problems are not limited to team studies but are often worse for studies of teams than individuals.

teams can reduce the fidelity of the experimental situation to real conditions, especially if rich communications are part of the real-world task setting.

5.5.2 Framework for Designing Team Research Methods

Cannon-Bowers and colleagues (Cannon-Bowers & Salas, 1997; Cannon-Bowers et al., 1992) have proposed a framework for conceptualizing and developing team measures. The purpose is to establish requirements for performance measures that are to be used in team training. Although developed in this context, the framework seems applicable to research uses as well.

The framework analyses potential measures in terms of two factors. The first is whether the measure assesses a team process or outcome. The second is whether the measure assesses a team characteristic or an individual characteristic.

The framework is also based on several assumptions. First, it is assumed that task and situational demands determine the need for teamwork to accomplish the task. Thus, teamwork is a response to the situation, which determines the specific kinds of interactions between teammates. Finally, it is assumed that outcomes can range from attitudinal to organizational level factors. In other words, results of teamwork can be measured not only in terms of accomplishment of the task but also in terms of the attitudes members develop and the cohesion of the team.

A 2x2 classification of measurement factors is presented in Table 5.9 (see Cannon-Bowers & Salas, 1997). The factors indicate the kinds of factors that can be assessed according to the combinations of the two factors above. They serve as a conceptual guide to researchers for selection of appropriate measures to address experimental issues.

	Process	**Outcome**
Team	Shared mental models Cue/strategy associations Task organization Compensatory behaviour Collective efficacy Dynamic reallocation of function Task interaction	Mission/goal accomplishment Aggregate latency (response time) Error propagation Mission-related error Aggregate accuracy
Individual	Assertiveness Information exchange Task-specific role responsibilities Procedures for task accomplishment Cue/strategy association Mutual performance monitoring Flexibility	Accuracy Latency Errors Safety Timeliness Decision biases

Table 5.9 – Classification of Measurement Factors (from Cannon-Bowers & Salas, 1997)

Team outcome measures assess the effectiveness of the team in accomplishing its goals. Team process measures, in contrast, identify how the team as a unit accomplishes task-related objectives. Individual outcome measures are used to determine whether team members can demonstrate needed team-related KSAs (i.e. function in teamwork as well as taskwork). Individual process measures are used to indicate the ways in which members accomplish tasks and how they contribute to overall team functioning.

Suggested measurement approaches are provided in Table 5.10.

	Process	Outcome
Team	• Observational studies • Expert ratings • Content analysis • Protocol analysis	• Observational scales • Expert ratings • Critical incidents • Automated performance recording
Individual	• Decision analysis • Policy capturing • Protocol analysis • Observational scales	• Automated performance recording • Critical incident • Expert ratings • Archival records

Table 5.10 – Suggested Measurement Approaches (from Cannon-Bowers & Salas, 1997)

5.5.3 Baker and Salas' Principles of Team Measurement

Although Cannon-Bowers and Salas' (1997) classification identifies key team measurement factors and suggests some general techniques, there remains a gulf between theory and practice. In particular, researchers must address the three questions of what, when, and how discussed above (Baker & Salas, 1997). Baker and Salas (1992, 1997) developed a set of principles for designing team measures, which are listed in Table 5.11. These principles provide guidelines for researchers developing measures for specific team studies.

5.5.4 Methods

This section reviews some specific team research methods. These methods are not meant to form a comprehensive set but, rather, to illustrate several team measurement approaches. Specific measures should be developed and/or adapted for the specific purposes of the given study. Elliot et al. (1996) provide lists of team measures for AWACS crews in several areas (see Annex A): *outcome measures* (overall measures of the degree to which teams accomplish mission goals), *indices of team SA* (degree to which members maintain current and accurate awareness and attentional focus), and *communications measures* (amount and type of communication). Researchers may want to consult these lists when designing empirical studies.

Principles of Team Measurement Design
Principle 1) To understand teamwork, researchers need a good theory
Principle 1a) Fully understanding team performance requires behavioural, cognitive, and attitudinal measures
Principle 1b) Team theory has developed fast but there needs to be greater empirical work to develop team measures
Principle 2) Researchers need multiple measures at different times in team history
Principle 2a) Measures must capture dynamic aspects of teamwork and assess the effects of developmental processes on team measures
Principle 2b) Measures must reflect the maturation process of the team
Principle 2c) Measures must account for team member experience
Principle 3) Measures should be observational and infer team KSAs from behaviour
Principle 3a) Team performance is not simply represented by what members do
Principle 3b) Observation is critical for measuring behavioural skills
Principle 3c) Measures that assess shared mental models and interpositional knowledge must be developed and validated
Principle 4) Researchers need to develop measures in a wide variety of contexts and tasks
Principle 4a) Researchers must develop measures of team KSAs
Principle 5) Measures must be reliable
Principle 5a) Reliability studies must reflect the characteristics of the measurement tool (i.e., understand the unique aspects of the measurement tool and how they affect the measure)
Principle 5b) Performance raters must exhibit high levels of agreement (90%)
Principle 5c) Team performance measures must demonstrate internal consistency
Principle 5d) Researchers need to establish reliability of team performance itself because teams may act differently at different times (maturation, situational effects)
Principle 6) Measures must be validated for practice and theory
Principle 6a) Content and construct validity of team performance measures must be determined
Principle 6b) Valid team performance measures must contribute to the development of valid team performance theories
Principle 6c) Criterion-related validity of team performance measures must be determined
Principle 6d) Team performance measures must predict team outcomes
Principle 6e) Team performance measures must look like they assess team performance (i.e., have face validity)

Table 5.11 - Principles for Designing Team Measures (from Baker & Salas, 1992, 1997)

5.5.4.1 Critical Team Behaviours Form (CTBF)

This is a checklist developed by Morgan et al. (1986, cited in Baker & Salas, 1992) to assess teamwork behaviours in seven areas:

- Cooperation.
- Communication.

- Team spirit or morale.
- Giving suggestions or criticisms.
- Adaptability.
- Coordination.
- Acceptance of suggestions/criticisms.

The CTBF relies on SME raters to observe teams performing in real or simulated work situations and record instances of critical team behaviours in the seven areas. The CTBF has been used in several studies and scores on the measure do seem to correlated with overall team performance (Baker & Salas, 1992).

The CTBF is only one type of rating scale developed for team measurement and many studies have employed other versions of the rating method (e.g., Adelman, 1992; Cyrus, 1991; Morrison et al., in press). Brannick et al. (1995) evaluated the validity of rating measures of teamwork. They had expert raters observe US Navy air teams as they worked through scenarios. The raters assessed teams on dimensions of measurement used by the US Navy Air Crew program (assertiveness, decision making, adaptability, SA, leadership, and communication). Brannick et al. (1995) found good convergent and discriminant validity of the rating measures and means ratings of all six team processes correlated with overall measures of mission effectiveness. They concluded that independent raters can provide sound assessments of team performance. One issue of concern, however, was that ratings of teams were not always consistent across scenarios, suggesting that raters are context sensitive. For this reason, Brannick et al. (1995) recommend using other measurement techniques with ratings.

5.5.4.2 Targeted Acceptable Responses to Generated Events or Tasks (TARGETS)

TARGETS is a scenario presentation technique (see Fowlkes et al., 1994) that employs rating techniques to assess team processes and performance. Thus, TARGETS illustrates how rating techniques can be used in the context of simulation.

The first step in TARGETS is to identify team skills or the behaviours that could occur in an operational setting (these will form the basis of the rating checklists). The second step is to determine which team behaviours can occur at any given critical point in a scenario. Scenarios are designed to provide measurement opportunities for all team behaviours in operationally relevant routine and critical contexts. Acceptable task responses are defined a priori by task analysis, review of standard operating procedures, and/or SME judgments.

Independent raters are given checklists based on this analysis and trained to detect and code relevant team behaviours. During scenarios, raters observe and record how the team performs. The checklist can be revised and refined to allow raters to detect relevant behaviours. This technique has yielded ratings with good validity and reliability (Fowlkes et al., 1994).

5.5.4.3 Team Interactive Decision Exercise for Teams Incorporating Distributed Expertise (TIDE^2)

TIDE^2 is another scenario presentation technique that can be combined with rating or observational measures, such as quantity of work, communications throughput, and

mission accomplishment (Weaver et al., 1995). It focuses on assessment of team decision making in anti-air teams (Ilgen et al., 1993). Teams consist of four participants, one a leader and three subordinates and work at low-fidelity workstations when they attempt to identify incoming targets and their intent. An advantage of $TIDE^2$ is that it can accommodate a wide range of scenarios and team factors. In addition, the workstation displays can be reconfigured in many ways to study the effects of numerous system variables. Thus, it is suitable for studying the impact of many team factors, such as hierarchical structure, uncertainty, leadership, and communication (Weaver et al., 1995).

5.5.4.4 Team Performance Assessment Battery (TPAB)

TPAB is a more generic scenario presentation technique in which teams work on realistic problems using low-fidelity workstations (Weaver et al., 1995). An advantage of TPAB is that it permits multiple tasks, including monitoring tasks, to be imposed on teams to assess factors related to workload. Thus, TPAB is suitable for assessing resource management aspects of team performance. Specific measures can be rating-based or performance-based (reaction time, accuracy, etc.).

5.5.4.5 Shared Mental Models

Team SA is an important aspect of teamwork but generally cannot be directly assessed by raters in the same way as behavioural components of teamwork. Kraiger and Wenzel (1997) propose four techniques to assess team SA by measuring factors related to shared mental models among team members.

The first measure examines how members perceive and process information. Team members perform conceptual mapping tasks (Section 5.4.4), which maps the individual conceptual structures within the team. Researchers determine the extent of agreement and the areas of overlap in knowledge and relational structure. This measure assumes that team SA consists of shared knowledge and does not account for the placement of individual KSAs in specific roles and responsibilities.

The second technique employs similarity ratings of team members to determine the conceptual organization of knowledge (Goldsmith & Kraiger, 1997, cited in Kraiger & Wenzel, 1997). Similar to the first technique, team SA is assessed as the degree of overlap in conceptual maps of individual team members.

The third technique infers team SA by the cohesiveness of the team. This measure assumes that shared attitudes is a component of team SA. Individual attitudes are assessed by questionnaire or interview methods. Informally, team SA is inferred by the degree to which team members' attitudes about the team overlap and the degree to which members can predict each other's attitudes. Kraiger and Wenzel (1997) note the need to develop formal methods to determine the overlap of individuals' attitudes.

The final technique to assess team SA is to measure the degree to which team members can predict what other members will do at any point of a simulation. Thus, as teams perform a scenario, members are queried at intervals about the knowledge states and actions of the other team members. These judgments are compared against self-assessments of knowledge and objective measures of behaviour for the team members. Team SA is assessed by the extent to which predictions match these measures. This technique captures the view of team SA that each member has knowledge needed for

his or her role and assesses the extent to which each member is doing what he or she is supposed to do.

Further work to develop more sophisticated measures of team SA is needed. Elliot et al. (1996) argue that team SA should be measured as a weighted average or sum of knowledge held by team members. Unfortunately, there currently seems to be no agreed upon way to do this.

5.5.4.6 *Situation Awareness Global Assessment Technique (SAGAT)*

SAGAT was developed by Endsley (1987, 1988, cited in Endsley, 1995b) to measure SA at all three levels. It is based on the *freeze technique*, in which a simulation is periodically stopped so that participants can be queried about their perceptions and understanding of the situation. Stopping the simulation prevents the measurement device from interfering with performance in the simulation while assessing current SA, rather than collecting participants' memory for SA at the end of the scenario.

The specific queries used in SAGAT depend on the domain. To use this method, researchers must define SA requirements and design queries to that elicit levels of SA from participants. The query format, however, is general. Computerized versions of SAGAT have been developed for air-to-air tactical aircraft and advanced bomber aircraft (Endsley, 1990c, 1989a; cited in Endsley, 1995b), which could serve as models for development of a naval C2 version.

5.5.4.7 *Team Workload*

Another factor crucial to team performance is team workload. Increases in either taskwork or teamwork will increase workload and potentially impair performance (Bowers et al., 1997). Team workload can be assessed in a number of ways (Bowers et al., 1997):

Subjective Techniques. These consist primarily of self ratings or observer ratings of workload. They are easy to use but affected by factors other than workload, such as stress.

Physiological Techniques. Workload can be assessed by measures such as EEG that measure bodily reactions to task demands. These techniques, however, tend to be insensitive and are expensive and hard to use, especially in operational settings.

Dual Task Techniques. A common way to measure workload is to give teams a secondary task. The secondary task should be simple and it should be easy to measure performance. Increases in workload associated with the primary task are assessed by decreases in performance of the secondary task. This technique tends to be sensitive but itself imposes additional workload and can alter how teams perform the primary task.

5.5.5 Recommendations

- Promote R&D to develop measures of team SA and shared mental models, especially a means to assess the overlap of knowledge among team members.
- Make a commitment to evaluate teamwork and team performance (i.e., devote the necessary resources of time, personnel, effort).

- Conduct further surveys of team methods currently available and develop more observational and objective measures to reduce dependence on subjective techniques.

5.6 Human-Computer Interaction

In the highly computerized domain of C2, HCI is a crucial part of the design of DSSs and C2 systems. HCI is its own discipline (e.g., Norman & Draper, 1986, cited in Howell et al., 1993) but, so far, unequivocal answers have been limited. Most success in the field has come in developing guidelines for narrowly defined tasks common to all computer users (e.g., text editing) (Howell et al., 1993). Thus, there are few absolute standards and researchers and designers must infer specific HCI components on a case by case basis. In general, designers of C2 systems have several tools for resolving HCI issues (Howell et al., 1993):

- Growing scientific understanding of human information processing.
- Improving understanding of C2 as a control process.
- Increasingly useful collection of guidelines for designing HCI.

5.6.1 HCI Design Methods

To a large extent, HCI design is part of the overall design process discussed in Section 5.1. We will not repeat that discussion except to emphasize the role of rapid prototyping in HCI design. Commercial HCI guidelines generally recommend several cycles of the design process. Table 5.12 lists design steps provided on a US government web site that illustrate the general approach (http://cic13.lanl.gov/guiguide/Guiguid4.htm).

HCI Design Procedure
1. Project leader defines duration of, iteration, and schedules of user interface design meeting
2. Team members meet to: • Review and complete task list and usability goals • Identify primary application objects • Define by user type the actions to be carried out or application objects • Review preliminary design ideas • Design flow of control (menus, command) • Design objects (windows, dialogue boxes)
3. Project manager creates first draft of prototype and first draft of usability test plan
4. Team reviews prototype and test plan; revise
5. Project leader reviews and revises
6. Project leader presents prototype and usability plan to user representatives for comment and approval
7. Participants review prototype and confirm design
8. Project leader presents initial design
9. Iterate steps 1-7 to move prototype closer and closer to acceptable design

Table 5.12 – A General HCI Design Process (from http://cic13.lanl.gov/guiguide/Guiguid4.htm)

The HCI design process begins with a clear specification of the user's needs. Howell et al. (1993) list a number of cognitive and behavioural issues arising in HCI design (see Table 5.13). These focus on the cognitive operation of the user and how to maximize the system's effectiveness to the user. In particular, the designer must identify interface components that allow the user to exert ample supervisory control over what the system is doing (Howell et al., 1993).

Behavioural Questions Arising in System Development
System Planning
1. (Assuming a predecessor system) What changes in the new system require changes in numbers and types of personnel employed in the previous system?
2. What changes in tasks to be performed will require changes in personnel, selection, training, and system operation?
Preliminary Design
3. Of the various design alternatives, which is the most effective from the standpoint of behavioural performance?
4. Given the system configuration, will system personnel be able to perform all required functions effectively?
5. Will personnel encounter excessive workload?
6. What factors are responsible for potential error and can these be eliminated?
Detail Design
7. Which is the better of two or more subsystem/component design alternatives?
8. What level of personnel performance can one achieve and does this level satisfy system requirements?
9. What training should be provided to personnel?
10. Are equipment design and job procedures properly human engineered?
Production
11. Since the questions raised in this phase are primarily the concern of industrial engineering, they are not discussed.
Test and Evaluation
12. Have all system dimensions affected by behaviour variables been properly human engineered?
13. Will system personnel be able to do their jobs effectively?
14. Does the system satisfy its personnel requirements?
15. What design inadequacies exist that must be rectified?
System Operations
16. Do any behavioural problems still exist?
17. What is the specific cause of these problems and what solutions can be recommended?

Table 5.13 – Behavioural Issues of HCI Design (from Meister, 1985, cited in Howell et al., 1993)

The design process should make extensive use of prototyping techniques, such as storyboarding (Section 5.1.1.3; Manning, 1991; Miller et al., 1992). The prototype serves as a model of thee system interface and provides most, if not all, graphical input and output functions so that users can interact with a concrete interface. Through storyboarding, designers implement a model of the interface without actually building an expensive system.

5.6.2 Interface Measures

Webb et al. (1997) describe a number of HCI issues central to system evaluation. This is not a comprehensive set but illustrates some areas of C2 systems in particular need of evaluation.

Navigation. This refers to issues of the structure of information in the user interface and the ease of movement through that information. Evaluations include *momentum* (ability to proceed through information at a constant rate with minimal effort), *intuitiveness* (ability for a relatively untrained user to obtain desired information using the interface), *menu logic* (ability of the user to locate a desired command), and *display suite structure* (the consistency of the layout of the display with users/ expectations).

Display Design. This issue concerns the configuration and information content of the user interface. Evaluations include *display density* (amount of information present on the display), *target sizes* (size of targets and other symbols), *colour* (the compatibility of colour with the expectations of users), *text* (the appropriateness of fonts, sizes, and highlights), and *symbology* (identifiability and consistency of symbol usage).

Database Organization/Information Search. This issue concerns the capability of the user to query an active database for information. Evaluations include *query interface* (ease of use) and *database structures* (degree to which it supports users' models of the system).

Interface Hardware. This issue pertains to the physical aspects of the interface with which the user interacts. Evaluations include *keyboards*, *cursor control devices*, and *display characteristics*.

5.6.3 Recommendations

- Adopt a flexible interface evaluation strategy.
- Address behavioural questions during the design phase.
- Survey HCI methods of the commercial software industry to aid the development of an HCI evaluation strategy for the upgrade of the HALIFAX class.

6. Summary of Methodology in the Operational Context

6.1 General

6.1.1 Phases of Operations

The discussion of theoretical issues and empirical results in previous sections indicated the importance of four broad phases of naval operations:

- Context.[8]
- Mission Preparation.
- Watch Transfer
- On Watch (Surveillance, Threat Assessment, Threat Response).

The discussion also clearly indicated the need to learn more about the roles of context, mission preparation, and watch transfer on decision making and C2. Unfortunately, it appears that most methodologies have not been developed with these phases in mind. Existing methods and research tools are not necessarily inconsistent with the study of these phases but effort is needed to determine their applicability. The literature contains little in the way of systematic categorization or organization of research methods, at least as they relate to the naval C2 domain. This effort could take the form of development of methods to target a particular phase or development of comprehensive methods that examine decision making across all phases. In practice, some combination of both kinds of methods would likely prove most useful.

Currently, most research tools and facilities seem designed specifically for the study of implementation, and more specifically threat assessment and threat response. Laboratories and simulations typically deal with threat situations (air, surface, and subsurface) and vary mostly in the particulars of responses (i.e., whether examining a single decision maker or a team, how performance is assessed, etc.). A great deal of effort is often put into creating a faithful recreation of the tactical workstation, which can enhance the value of methods for study of implementation but divert attention from mission preparation and watch transfer. Without simulating the tools and resources available for these other phases, researchers will be at a loss to set up viable empirical studies of activities in these phases. A broadening of research facilities (laboratories, simulation platforms, etc.) is needed to enhance researchers' abilities to examine all phases of operations. A primary issue to consider is how much experimental control can be achieved in balance with a high level of fidelity to the naval command environment.

Similarly, the scenarios used in empirical studies and analysis techniques (e.g., CTA) should be broadened to cover mission preparation and watch transfer. Currently, scenarios tend to

[8] This is not strictly a phase but worth considering separately because of the influence of mission context on all aspects of operations.

deal only with threat situations (but see Webb & McLean, 1998) and present an individual or team with a specific tactical problem, such as how to respond to an air threat. These scenarios assume some level of preparation but do not make specific any steps that were taken, or could have been taken, to prepare for threat situations. Similarly, they tend to ignore previous experience and training as well. As we have seen (Section 4.2), however, preparation can play a significant role in how a decision maker ultimately responds to a situation. To further explore that connection, researchers must examine decision making across preparation and implementation phases.

6.1.2 Team Methodologies

Another general issue to consider is the need for team methodologies. A fair amount of work has been done, albeit in broader domains than naval C2, and a number of specific methods developed (Section 5.5.4). Nevertheless, some important issues remain, most notably what aspects of teamwork and team decision making should be measured (Baker & Salas, 1997). The theoretical ideas under examination will determine which behavioural or cognitive components should be assessed. Thus, researchers must be flexible in their approach and adapt methods to specific theoretical goals. In particular, researchers should be prepared to employ both process and outcome (Cannon-Bowers & Salas, 1997).

Another issue that needs to be resolved is the appropriateness of small teams to the study of naval C2. Driskell and Salas (1992) have argued that small teams can serve as models for larger teams and yield theoretically relevant, generalizable results. Researchers should be cautious, however, when attempting to study decision making of CTs. It may well be possible to meaningfully address issues of CT performance by examining artificially small teams, or even individuals, but only if the experimental setting realistically simulates the work domain and informational richness of real CTs and their relationships with other CTs within a taskgroup setting.

One area of richness to consider is the extensive communication within CTs, which helps team members establish and maintain SA. In addition, teams interact with other CTs within a task group. A danger of examining command performance in a small team setting is that the amount and detail of communication will be lost, affecting the performance of the participants. Communication is only one aspect of team functioning that could be affected by decreasing the size of the team. Role responsibility, cohesion, common intent, and coordination could all suffer if participants are examined outside the normal context of the full OR team (and perhaps outside the taskgroup setting).

Another crucial aspect of the CT is the command structure. Whether examining a small or full team, the research method should faithfully replicate a command structure. This is necessary to identify key roles, tasks, and lines of communication for team level decision making. Unfortunately, it may not be an easy matter to recreate this structure in an empirical investigation. Experimental studies will almost certainly involve artificial teams created solely for the purpose of the experiment. Participants may not even be active naval personnel. In this case, it is up to the researcher to identify the elements of command structure and replicate them in the design of the experiment. Even when using analysis techniques, researchers may not be able to examine the entire CT. In this case, it can be unclear what role other team members would play in real events or how their absence in the analysis alters the input of participants.

Finally, it is also important to remember that different teams can be involved at different stages of missions. The team that engages in planning may not be the same team that responds to a threat. Methodologies are needed to examine the interaction of teams within a vessel and within task groups (e.g., differences in mental models of the operational setting across teams). So far, the focus has been on studying how individuals within a team function together. This is inadequate to capture the full range of contexts in which CTs will function. Nor does it account for the watch transfer.

6.1.3 Team Situation Awareness

As mention earlier (Section 2.5.4), there is debate over the nature of team SA. One view holds that team SA consists of shared knowledge among all team members, whereas the other view holds that team SA consists of each member having knowledge appropriate to his or her role. Resolution of this debate is needed to develop methodologies to assess team SA.

Each view poses its own methodological problems. If we adopt the first view, we require a method to assess shared knowledge. Attempts have been made at developing such methods (Section 5.5.4.5) but these methods tend to rely on judgments of raters to quantify the overlap of team members' mental models. In contrast, if we adopt the second view of team SA, we no longer need worry about the extent of shared knowledge. We do, however, have to specify in detail the knowledge requirements of each team member. This requires a very sophisticated understanding of individual and team level tasks – an understanding that might not be possible prior to conducting empirical assessments.

In addition to methods for assessing team SA, researchers need methods for exploring demands on team SA associated with:

- Volume of information.
- Updating goals, ROEs, etc.
- Coordinating SA within the team and between other teams.

6.2 Operational Context

Perhaps one of the most difficult aspects of real operations to replicate in empirical studies is the mission context. There are a large number of variables that must be considered (e.g., location, disposition of friendly, neutral, and enemy forces, single ship versus task group operation, uni-national versus multinational operation, etc.) and anticipating operational context can be a challenge. As mentioned, one significant problem is recreating a realistic full CT in which to study decision making. In addition, current methods tend to omit consideration of other important contextual factors.

Scenarios and simulations, for example, typically present simplified missions and threats. A scenario may entail only a single threat at a time, even though the probability of multiple threats is high in real operational settings. Scenarios may also present unrealistically simple ROEs, ignoring such factors as the presence of neutral vessels, political implications, fatigue, stress, team cohesion, and so on.

One crucial factor that is often overlooked is that vessels will almost always serve as part of a task group and coordinate their activities with other ships and aircraft. Methodologies tend to involve only single ship activities, ignoring the interactions of CTs across vessels. This simplifies problems for CTs to some degree but presents participants with unrealistic missions. Given that task groups may

be multinational, it is probably also worth considering how CTs work with teams of other nations and other services as well.

6.2.1 System Design

It is not clear that current approaches to system design take into account the rich context in which systems are to operate. User-centred and prototyping methods tend to be task-oriented; that is, they focus on how the user performs predetermined tasks. This is important but the kinds of tasks performed by the CT and the nature of decision making will depend on the context. To be fair, it is unclear whether there can possibly be a systematic approach to modeling context given that the role of the system may change over time. Designers will probably be unable to predict all possible operational contexts. Nevertheless, consideration of the major underlying factors should improve system design.

There is a need to consider system design as broadly as possible so that designers can design systems to serve not just in the narrow confines of anticipated objectives but all possible objectives. The ecological approach to design (Vicente & Rasmussen, 1992) helps to accomplish this. The aim in this approach is to define the work domain, which consists of all the possible ways of accomplishing work objectives (in this case, mission objectives). The work domain is defined by the constraints on action that delimit how one can achieve the objectives. Consideration of the work domain prior to establishing system requirements allows designers to build into systems greater flexibility to achieve objectives in multiple ways. Defining the work domain also forces designers to identify contextual factors and determine how they constrain the work domain.

6.2.2 Analysis Techniques

Analysis techniques, such as CTA, are particularly dependent on context. How one performs tasks depends on whether one is in a team setting, operating in a task group or single ship, the nature of political implications, density of information streams, and so on. As noted above, most scenarios offer only a very impoverished context in which to analyze performance. To the extent that contextual factors are excluded, analysis techniques will fail to identify relevant strategies, decision requirements, and cognitive processes. Thus, it is essential that researchers develop detailed and realistic scenarios for use in analyses.

6.3 Mission Preparation

6.3.1 System Design

Just as current system design processes do not focus on context, they seem to generally ignore mission preparation as well. This may be due to a focus on technological systems and design of systems to respond to intense situations, such as threat response. Nevertheless, planning and preparation is crucial and should be explicitly included in the design process. It is during this phase that CTs create mental models upon which subsequent decision making depends.

In fact, mission preparation, watch transfer, and implementation should be considered as an integrated set of stages within the design process. Rather than view each as a problem area, it is better to view operations more holistically, with particular attention to the interactions between mission preparation and implementation. This view of design would emphasize the

integration of activities in planning and watch, perhaps improving the effectiveness of mission preparation.

6.3.2 Research Tools

Current laboratory facilities generally do not provide any environment in which to study mission preparation. In particular, they lack the means to bring together a team for mission planning and rehearsal with realistic tools and means to communicate pre-plans to a CT on watch. Similarly, scenarios typically do not address mission preparation.

6.4 Watch Transfer

6.4.1 System Design

The watch transfer has serious implications for CT readiness. A new team is suddenly put in place of the previous team and must rapidly "come up to speed" on the situation. This can be challenging due to the volume of information generated on the previous watch (messages, changes in friendly, neutral, and enemy forces, etc.) and the demands of establishing team SA. Consequently, it is important that watch transfer be actively considered in the design of C2 and decision support systems.

Some of the watch transfer issues that must be addressed are:

- Information processing demands associated with message traffic.
- Updating mental models of the tactical situation.
- Updating mental models of ROEs, operating procedures, pre-plans, and other mission related criteria.
- Achieving continuity of action with the CT of the previous watch.
- Achieving good team SA
- Coordinating with other CTs in a task group.

6.4.2 Research Tools

Research tools tend to ignore the watch transfer process. Simulations and scenarios are set up to investigate the performance of a team or individual *during* a watch. Any watch transfer activities are dealt with unsystematically in the instructions given to participants.

Examining watch transfer requires greater investment in facilities, time, and participation because researchers must employ at least two teams of participants. To accurately model watch transfer effects, the research design must follow one team as it deals with events on one watch then follow a second team as it deals with events on the following watch. Only in this way can researchers chart the differences in SA and decision making between teams as well as the difficulties the second team has in achieving SA. This would be especially cumbersome in experimental studies, which attempt to isolate a few independent variables for study. It would, however, be somewhat easier to accommodate in CTA and other analysis techniques, which track activities over undefined time periods.

6.5 On watch

6.5.1 Research Tools

Most research methods (and system design processes, for that matter) seem to deal with this phase in great detail. The focus of future work will not be on developing new methods but determining which methods are best suited to the naval tactical domain and how best to employ those methods.

To these ends, researchers should review methodologies in relation to the theoretical issues identified in this literature review. This literature review has highlighted a number of these issues, including:

- Determining the applicability of analytic versus intuitive decision making.
- Determining how decision makers achieve all three levels of SA.
- Determining how teams achieve team SA and coordinate team decision making.
- Link of decision making to previous planning and experience.

These are just a few issues and others are certainly relevant to C2 and decision support needs for the CPF.

In addition to theoretically relevant methods, however, researchers must also select or devise operationally relevant methods and measures. Real combat environments contain a number of variables that are extremely difficult to simulate in laboratory or even exercise settings. These include the stress, risk, and confusion associated with combat. The closer one gets to actual operations, the more applicable will be empirical results. Thus, methods that can be integrated with exercises and/or operations could prove very valuable to research on decision support for the CPF.

6.5.2 Teams

Teamwork during threat assessment and threat response can be difficult to study because the CT is engaged in a lot of activity with little time to handle additional tasks. This presents the twin challenges of a) selecting measures from the mass of potential data that could be collected and b) recreating realistic conditions that will promote the kinds of teamwork behaviours exhibited in actual operations. One empirical approach is to allow CTs to engage in team behaviours but periodically interrupt the team to apply measures. This deals with the first problem by creating sufficient time to thoroughly assess relevant variables but exacerbates the second problem by disrupting the team's activities. Another approach is to allow the CT to perform scenarios without interference, applying strictly observational measures. This minimizes the second problem but burdens the researcher with a large amount of data, making analysis difficult. Some compromise between these approaches is needed.

An aspect of teams that should not be overlooked is the experience that each member brings. In part, experience contributes to the role structure within the team, with each member performing specialized functions. However, differences in experience between members can potentially affect how team members interact. In particular, research is needed to determine how team members of different backgrounds understand one another. Does an ORO with previous experience as a SWC interact differently with the CT than an ORO with experience in communications? It is likely that in the first case the ORO would have better understanding of the SWC than other positions in the CT. The ORO in the second case may have difficulty

with other positions. Beyond simply affecting how team members communicate, past job experience could affect the kinds of decision making strategies that individuals bring to their teams.

6.5.3 Human-Computer Interaction

HCI issues are perhaps more critical during the watch than at other phases. Events during the watch can place the CT under severe time pressure, making an interface that promotes fast, accurate performance a critical component for success. During planning, the CT may have time to work around a clumsy interface but, when confronted with a threat, the CT team cannot spend time figuring out how to accomplish an action with an interface that is not intuitively clear. The on watch team especially needs an interface that provides immediate updating of changes in the situation to allow the CT to maintain good SA and a mental model of the tactical situation.

7. Guidelines

This section reviews guidelines developed for naval C2 systems and functions, as well as general guidelines for HCI. More work needs to be done to develop comprehensive guidelines specific to naval C2 systems. Currently, guidelines are not as readily available as they could be.

7.1 DSS Design Guidelines

Perhaps because the design of DSSs depends so much on the particular domain, there seems to have been less work on developing guidelines for design than guidelines for the design process itself (see Section 5.1). Where design guidelines for naval C2 exist, they tend to deal with high level design issues, such as specifying decision requirements or interface characteristics (which will be discussed in Section 7.3 below). Ramamurthi, King, and Premkuman (1992), for example, report a literature review of factors that contribute to the success of DSSs. They consider user, technology, task, and organization factors and model their effects. These factors can serve as a guide to issues that should be considered in the design of a DSS.

Areas of Management in C2
Information Management
Volume of information
Differing types of information to manage
Relative accuracies of reporting platforms
Accuracy of sensors
Timing delays with communications
Lack of unique track ID
Platform management
Control of ship movement
Integration of ship information with tactical information
Resource management
Management of limited resources (fuel, ammunition, etc.)
Monitoring weapons systems and other resources
Integrating resource management tools with tactical displays
Tactical Management
Control of assets
Tactical situation picture
Kinematic information
Intelligence information
Operational information from other commanders

Table 7.1 – C2 Management Functions in Need of Support (from Anderson, 1990)

A number of researchers have analyzed C2 to determine the major decisions that must be made (Anderson, 1990; Klein et al., 1997; Speed, 1994). Anderson (1990) divided C2 into four areas of management (discussed in Section 2.1). Within each area, Anderson (1990) specifies some of the specific factors contributing to the need for decision support (see Table 7.1). The Generic C3I workstation developed by Anderson (1990) also serves as a kind of guideline for development of tactical DSSs (see description in Section 3.2.7.3).

Speed (1994) provides an analysis of decision criteria to assist command level decision makers in evaluating their decisions. These criteria, listed in Table 7.2, focus on determining an optimal decision and may be more suitable to the strategic than tactical level of decision making. At the least, weighing multiple options is difficult under the severe time pressure of most tactical situations.

Procedural Criteria for Rational Decisions
Appraise the operational mission
Understand goals and purposes within strategic context
Clearly convey goals and purposes to staff
Commander should facilitate problem-solving by staff but let staff solve problems
Thoroughly canvas a wide range of alternative courses of action
Operational commander must be open to all possible courses
Provide atmosphere where all viewpoints/suggestions are welcome
Carefully weigh the costs and risks of each alternative
Designate someone to be a "critical evaluator" to ensure criticism of plans
Regulate the criticism process so prolonged debates and emotions do not impair morale or effectiveness
Implement, continually re-evaluate, and provide feedback for selected COA, modifying operational plan as necessary
Be open to change because the situation will never be completely clear
Continually obtain feedback about achievement of goals
Maintain flexibility to minimize effects of unforeseen problems or to move plan in a new direction

Table 7.2 – Decision Criteria for Tactical Decision Making (from Speed, 1994)

7.2 Measures of Effectiveness

7.2.1 DSS Evaluation

Guidelines have also been developed to assist in the specification of MOEs to evaluate DSSs and other C2 systems. Leroy and Hofmann (1996) developed a set of criteria (discussed in Section 3.1.1.2) to indicate the system attributes that should be assessed when measuring the effectiveness of a system. The criteria specify both system characteristics (e.g., survivability, computation capacity) and characteristics related to the use of the system (e.g., information quality, timeliness). Leroy and Hofmann (1996) also list desired characteristics of MOEs themselves.

7.2.2 Team Performance

Cannon-Bowers and Salas (1997) developed a framework for conceptualizing TMPs that can be used to assess the progress of teams in training or performance on the job. They are designed to measure both team and individual level performance and address processes as well as outcomes. Table 7.3 lists requirements for TPMs in training. Cannon-Bowers and Salas (1997) also developed a framework for developing TPMs for individual and team level process and outcomes, which was discussed in Section 5.5.2.

Requirements for Team Performance Measures (TMPs)
TMPs must consider multiple levels of measurement
Both individual and team-level
Team competencies exist at several levels
Team competencies held at the individual level; generic KSAs with respect to task or team involved
Team competencies held at team level; specific to team or task
TMPs must address process as well as outcome
Outcome measures are not diagnostic because they do not indicate the underlying causes of performance
Performance measures more directly describe performance of interest
TMPs must be able to describe, evaluate, and diagnose performance
Measures must be able to describe behaviour accurately
TMPs must be dynamic; adjust to changing situational characteristics, and be sensitive to small deviations in performance
Modern tasks require manipulation of advanced equipment, often have largely cognitive demands (no behavioural indicators); complicates issue of recording team performance
Observed behaviour must be assessed as to appropriateness given task situation; develop standards
Standards can be normative or criterion-based
Most important feature of TMPs is diagnosticity – ability to identify underlying causes
TMPs must provide a basis for remediation
Provide trainees with knowledge of results
Feedback that indicates how performance can be improved
Drive selection of subsequent instruction
Aid in structuring content of training
Feedback must be accurate and timely
TMP used to determine level of material appropriate to trainee as well as preferred learning style

Table 7.3 – Training Performance Measure Requirements (from Cannon-Bowers & Salas, 1997)

7.3 Training

Numerous books and articles deal with training issues in general, such as the use of multimedia or interactive instructional techniques (Bostrow et al., 1995; Miserandino, 1998). We have already discussed two issues related to training for teams.

7.3.1 Team Competencies

Cannon-Bowers et al. (1995) developed a classification of team competencies necessary for effective team performance, discussed in Section 2.5.1.2. This framework can be used to specify elements of team training programs. In particular, training should consider two factors:

- Whether competencies are specific to the particular team.
- Whether competencies are specific to the particular task.

Table 2.3 presents a breakdown of team competencies and Table 2.5 a list of propositions for effective team training.

7.3.2 Team Requirements

Swezey and Salas (1992) developed guidelines for team training programs. Training needs depend to a great extent on the task, organization, and context but these guidelines can serve as a starting point and help identify areas where further development are needed.

Swezey and Salas (1987, cited in Swezey & Salas, 1992) classified team training guidelines by 12 categories that address a team process principle:

- Team mission and goals.
- Environment and operating situation.
- Organization, size, and interaction.
- Motivation, attitudes, cohesion.
- Leadership.
- Communication.
- Adaptability.
- Knowledge and skill development.
- Coordination and cooperation.
- Evaluation.
- Team-training situation; general, role of instructor, training methods.
- Assessments of team-training programs.

Swezey and Salas (1992) surveyed 2000 documents to locate guidelines pertaining to these categories. They located 150 articles containing such guidelines and extracted over 500 specific guides, which are presented in their article. The guidelines are specific and cover all the categories established by the authors. They should prove helpful in the development, evaluation, and revision of training programs. One weakness of these guidelines, however, is that they are aimed at traditional instruction techniques and offer few direct guides on multimedia or experiential learning.

7.4 HCI and Visual Displays

HCI is perhaps the area where the most work has been done to develop practical guidelines, both in general and specifically for C2 (see Webb et al., 1993).

7.4.1 Commercial Guidelines

The computer software industry has developed extensive guidelines for HCI, many of which are freely available through the WWW. These guidelines tend to deal with standard computer workstation applications but the rules for display design and information presentation can be readily adapted to naval C2 applications.

IBM (www.ibm.com/ibm/hci/designer/uiarchnt.html), for example, has produced guidelines for user interface design. An advantageous feature of these guidelines is that they contain information at several levels, indicating IBM's design principles (statements of policy), design guidelines (functional features that should be used), and design conventions (specific implementation features and processes).

Additional guidelines are available from Microsoft, Motif, Tandem, and Bellcore, among others (www.dordt.edu:457/VYCLG/vtclgD.style_standards.html).

7.4.2 C3I Interface Guideline

Howell et al. (1993) have developed visual display principles specifically for C3I systems. These guidelines are intended to assist designers in during the conceptual stage of system development. The display principles are based on the major task categories of C3I (see Wohl, 1981, cited in Howell et al., 1993) and a set of behavioural questions pertaining to system development (see Meister, 1985, cited in Howell et al., 1993). The classification of operations is shown in Table 7.4. These guidelines provide definitions at each level in the classification and specify design principles. These are not specific implementation practices but functional considerations that should be achieved by the designers. They should help the designers identify the importance of specific actions, information needs, and computations.

7.4.3 SA Guidelines

Several researchers have considered the implications of HCI and display design for SA. These researchers have not developed extensive guidelines but do offer some specific recommendations about how to design displays to facilitate SA (see Section 2.3.4). In particular, Endsley's (1995a) taxonomy of SA serves as a major guide to design of support.

Garner and Assenmacher (1997) have developed even more comprehensive SA guidelines. The purpose of their guidelines is to assist designers identify requirements for SA and ensure that systems meet these requirements. The guidelines consist of basic HF issues that can be addressed before task-related issues are considered. The guidelines should be consulted to create checklists of design features for interfaces aboard the CPF.

Garner and Assenmacher's (1997) guidelines are organized into 7 functional areas:

- System features.
- Display formatting.
- Information coding.
- Enhancement coding.

- Auditory coding.
- Environmental stressors.
- Advanced technologies.

Within each area, Garner and Assenmacher (1997) identify factors that could affect SA. They serve to indicate issues that should be specifically addressed by design features of a system.

Major Operations Performed on Information in C3I Tasks	
Simple Extraction	
1.0	Read-out
2.0	Identify/recognize
3.0	Locate
Complex Extraction	
4.0	Discriminate/compare
5.0	Filter/ignore
6.0	Perceive/interpret pattern
7.0	Correlate
8.0	Monitor
Process	
9.0	Remember
10.0	Estimate
11.0	Calculate
12.0	Integrate/organize/aggregate
13.0	Evaluate
14.0	Generate/create
15.0	Choose/decide
16.0	Manipulate
17.0	Command system
Multiple Operations	
18.0	Complex interaction

Table 7.4 – Classification of C3I Task Functions (from Howell et al., 1993)

7.5 Recommendations

- Promote R&D to develop guidelines in all areas pertaining to C2 and decision support.
- Create an integrated set of guidelines covering all major aspects of decision making, SA, teamwork, and training for the CPF by conducting a literature survey to identify existing guidelines and obtaining commercial HCI guidelines.

8. Summary of Guidelines in the Operational Context

The section on guidelines is not well developed. In part, this reflects limits on the time and effort that we could spend locating and describing guidelines in this literature review. However, in comparison to research on theory and methodologies, less effort has been devoted to the development of guidelines specifically for naval C2 and decision support.

Although the guidelines section is brief and incomplete, a few general lessons did emerge from it. First, there appears to have been no systematic effort to develop guidelines for all aspects of C2 and decision making (e.g., teamwork, training) and all phases of operations (in particular, mission preparation and watch transfer). Thus, existing guidelines tend to be haphazard and only partially relevant to the upgrade of the HALIFAX class. Guidelines in some areas, such as HCI, are extensive and immediately applicable but guidelines in other areas are only minimally developed (SA, cognitive fit in interface design, etc.). A separate literature review to identify and integrate guidelines would be valuable to the upgrade of the HALIFAX class. The goal should be to develop a comprehensive set of guidelines selected specifically for the HALIFAX class, addressing all aspects of C2 and decision making.

Second, the guidelines identified in this literature review deal almost exclusively with implementation issues (surveillance, threat assessment, and threat response). In some cases, this problem is less pronounced. HCI guidelines, for example, have been applied in the context of workstation design for the OR room but are fairly general. With little effort, these could be extended to tools for preparation and watch transfer. In other cases, such as guidelines for SA, no consideration has been given to mission preparation. Development of SA guidelines for this phase will probably require greater effort and empirical investigation. There may not currently exist sufficient guidelines for all areas relevant to the design of decision support for the CPF.

9. Summary and Conclusions

This section summarizes the major issues and lessons learned.

9.1 Approaches to C2

9.1.1 Definitional Issues

A major finding of this literature review was that work remains to be done on the basic definitional issues of C2. This is not a simple concept and there seem to be many views on C2 within the scientific and military communities. The lack of a single consistent definition could have far-reaching affects for R&D of decision support for the HALIFAX upgrade because it will impair efforts to specify clear objectives for future decision support. Furthermore, the interpretation of existing research can be problematic because it is often unclear what definition of C2 underlies that work.

One aspect of this definitional confusion centers on the relative weight assigned to the command versus control aspects of C2. Pigeau and McCann (1995) have pointed out that the two terms are not equivalent but, rather, describe separate contributing components of C2. Nevertheless, much the research on C2 has focused on just one aspect, namely control. Control deals with the implementation of authority and how direction of forces can be accomplished. By focusing on this aspect, researchers and military planners have ended up adopting a technology-driven, process-oriented approach to C2. Commanders are seen as managing a system and so the goal of C2 and decision support becomes to improve the commander's ability to know what is going on and effect changes in the system.

In contrast to this mechanistic approach, command deals with the authority of individuals to direct forces and the way legal and military responsibility is allocated and enforced. In particular, command deals with the way intent, either explicit or implied, is conveyed by commanders to subordinates and the extent to which subordinates are able to act in order to accomplish that intent (in contrast to merely fulfilling an established process). Thus, the study of command entails greater emphasis on military organization and team function.

This literature review reflects the current prevailing emphasis on control. As a result, it has little to say about issues of intent, authority, and military organization. This does not mean, however, that these areas are secondary to control issues. Rather, more research is needed to study command. It is crucial to determine well in advance of any system development what kind of C2 will be adopted because this definition will determine what will be needed for the HALIFAX upgrade and what resources will be available.

9.1.2 The Future C2 Environment

It is widely anticipated that NATO naval forces will be called upon to serve frequently in littoral environments in the future (e.g., O'Neill, 1996; Sundin, 1996). This will create a large number of problems for effective C2 (e.g., greater uncertainty, greater constraints on movement and action). Thus, another conclusion of this review is that R&D efforts should be directed at addressing the specific issues raised by littoral operations. In particular, R&D

must address ways to deal with the tremendous complexity of these situations and the stress that it induces in crews.

Although it is prudent to plan for a shift to littoral operations, however, it must be remembered that this is merely a projection of future needs. Perhaps a chief need for future C2 systems will be flexibility – flexibility to deal with unanticipated situations and events and flexibility to be modified and upgraded quickly and efficiently if the need arises.

9.2 Theoretical Issues

9.2.1 Decision Making

Review of the literature regarding decision making and C2 revealed that there are two major competing perspectives, analytic and intuitive. The analytic perspective comes from traditional research in AI and the behavioural sciences and is primarily an attempt to formalize the process of decision making. The intuitive perspective applies a descriptive approach to determine how expert decision makers solve problems within a domain.

The analytic and intuitive approaches differ in many respects. Analytic theories are highly prescriptive, general, and optimizing, whereas intuitive theories are highly descriptive, domain-specific, and satisficing. Thus, each has distinct strengths and weaknesses. It seems unlikely that an entirely analytic approach or an entirely intuitive approach will serve as a useful model for every aspect of naval tactical decision making. Consequently, effort should be directed at employing the advantageous aspects of each approach; i.e., identifying the areas of C2 to which each perspective is applicable and synthesizing the two approaches in a theory specific to naval C2.

To this end, there are some useful distinctions within decision making and C2. First, decision making occurs at a range of levels within the military organization. Strategic, tactical, and procedural levels of C2 serve somewhat different functions under different constraints (Flin, 1998). Thus, it is important to determine how analytic and intuitive approaches apply at these levels. Second, within the tactical level, C2 involves a number of different activities. An often-overlooked part is planning (but see Webb & McLean, 1997). Crews may spend little time actually dealing with a critical event (e.g., potentially hostile aircraft) compared to the time spent anticipating events and planning responses. Most research, however, addresses decision making in critical events and ignores planning. This focus has lead, perhaps, to a misrepresentation of what tactical decision makers do and, hence, their decision support requirements.

9.2.2 Situation Awareness

SA is a critical concept for C2 and decision making, especially within the intuitive approach. Despite this fact, there appears to be only one major theory of SA, that proposed by Endsley (1995a; 1997). Endsley's theory is useful from a definitional point of view because it distinguishes three levels of SA and provides a framework for identifying the factors affecting SA and causes of errors in SA. The theory, however, does not describe in great detail the processes by which people form SA, making it difficult to fully integrate SA into theories of decision making.

Perhaps the major implication of Endsley's theory of SA is that a person's awareness and understanding of the situation has more to do with how information is gathered and presented than with the amount of information that is available. Thus, access to a vast amount of information does not guarantee good SA. In fact, it may actively impair SA. Because SA involves creating a mental model of the situation and linking it to one's goals, good SA arises when a person can effectively process data about the situation and relate it to goals and schemata for previously encountered events. This implies that less data can be better, as long as it conveys the important features of the situation. In terms of designing effective decision support, Endsley's theory points to a need to emphasize the ease of data extraction and the ability to obtain task- and goal-related information with minimal effort.

9.2.3 Teams

The most important aspect of team tasks is that they impose two kinds of functions on the individuals making up a team, taskwork and teamwork. The task, of course, dictates goals and processes that must be accomplished but in addition team members must engage in a number of activities aimed at making the team itself work properly (e.g., coordination, communication). Some research has addressed the nature of teamwork but it seems that more is needed, especially to establish the nature of teamwork in the naval C2 context.

One approach has been to study teams and identify the basic competencies or KSAs that mark good teams. Cannon-Bowers et al. (1995) have established a classification scheme that breaks team competencies into task-specific versus task-generic and team-specific versus team-generic. This scheme points to different areas for further research, specifically to identify the behaviours and skills that indicate good teamwork in general and the behaviours and skills needed specifically in the naval C2 domain.

One concept that emerges as a key to good teamwork is shared SA. Although some debate still rages concerning exactly how to define shared SA, coordination, communication, and most other teamwork activities rely on team members having knowledge in common. In particular, members need to understand the responsibilities and duties of their teammates in order to support them (i.e., with needed information, resources, etc.). This is one necessary component of form common intent within the team (see Pigeau & McCann, 1998). The crucial idea behind this is that a team functions as a support system for each member at the same time that members cooperate to accomplish the team goals. This point is, perhaps, disguised in some team research that focuses in detail on the specific team processes that govern behaviour. It does, however, point out the need to integrate the team concept in the design of decision support.

9.2.4 Expertise and Training

The development of expertise is critical to effective C2 – so much so that expertise should be considered a part of decision support as much as a goal of training. Expertise has two important features. First, it is stage-like in the sense that people not only exhibit different levels of performance as they acquire expertise in a domain, they also exhibit qualitatively different strategies for performing (Anderson, 1995). Thus, expertise is more than an accumulation of knowledge and skills (although this is important as well). It has to be viewed as the periodic reorganization of knowledge and skills to produce a more effective performance system. Second, expertise depends on organizing acquired knowledge in a way

that facilitates effective problem solving. Thus, experts organize their domain knowledge in terms of the goals, functions, and constraints of the domain, which allows them to recognize problem types and apply the appropriate solutions.

The nature of expertise implies the need for support of the transitions between stages of expertise. Some combination of training and onboard DSSs must help C2 team members move from an initial declarative understanding of the domain to an automatic and procedural level of performance.

Modern training techniques should facilitate this kind of support very well. There has been a great deal of progress in the development of experiential learning techniques that help people gain expertise in a domain by actively solving problems. These techniques focus on learning-by-doing and engaging in real-domain tasks and are supported by interactive instruction and feedback that allow a trainee to ask questions and get answers immediately. Multimedia techniques are also useful for implementing this style of instruction. Although traditional lecture-based training will have its place, especially early on when complete novices will require declarative knowledge to form an initial understanding, learning-by-doing will promote faster learning and better performance.

9.2.5 HCI

The literature review has identified one main principle of HCI that should be adhered to in the design of any C2 system, namely cognitive fit. This is the principle that a system should represent information is a form directly usable by the human operator. In other words, there should be a fit between the way the operator thinks about his or her tasks, processes information, and issue commands and the way the system displays information and gathers input from the operator.

Cognitive fit must be based on a thorough understanding of the domain and the tasks of the operator. Ecological Interface Design (Vicente & Rasmussen, 1992) is one approach to defining the user requirements for a system. It is particularly useful in that it focuses not just on the layout of the computer display and input devices but also on the conceptual design of the system. That is, this approach attempts to define design elements in terms of their ability to support a user's skills and knowledge.

9.3 Empirical Results and Lessons Learned

A comprehensive review of the vast empirical work on decision making and related issues was beyond the scope of this literature review. Instead, it focused on a few important issues that could affect development of decision support and training for the HALIFAX class.

9.3.1 Decision Making

The major finding in this area is that tactical decision makers overwhelmingly employ intuitive decision making strategies. Although the precise breakdown of recognition, story building, and other intuitive strategies is not clear, decision makers rely on memory and experience to deal with tactical problems. Thus, a simple characterization of expert problem solving in the naval tactical domain is:

1. Identify that a problem exists.
2. Gather data to characterize the situation.
3. Recognize and classify the situation.
4. Retrieve a workable COA.
5. Verify that the COA will work.

The first two steps broadly consist of generating SA and will have to be iterated many times before proceeding. The third step involves matching the perceived situation to a mental store of events and problems. The fourth step relies on simple associative memory to call to mind COAs that have worked in the past. The final step consists of mental simulation to predict the consequences of implementing the COA.

Overall, this kind of strategy is used to deal with the uncertainty and severe time pressure of tactical situations. This implies that tactical decision makers are in primary need of support for SA, recognition, and mental simulation.

Before accepting this implication, however, we must remember that most research on naval C2 has considered only crisis situations, such as dealing with a potential threat. Research typically ignores mission planning and other activities, although Webb and McLean (1997) documented extensive planning activities of OR teams aboard the CPF. Planning is typically not done under severe time pressure, although data may be uncertain. One issue that warrants exploration is whether analytic approaches to decision making can offer a better account of planning activities than intuitive approaches.

9.3.2 DSSs

A great deal of effort has already been directed at developing DSSs for naval tactical decision making and C2. There are three broad approaches: ESs, interface DSSs, and intuitive DSSs. ESs represent, perhaps, a traditional approach in which a computer system computes recommended COAs for a human user. In contrast, both the interface and intuitive DSSs attempt to improve the efficiency of human decision making rather than perform the decision making itself. Interface DSSs do this by organizing data in a form that is most readily usable by the human decision maker. They perform only simple computations aimed at processing raw data into a more useful form. Intuitive DSSs take on more computation for the user and perform some decision making steps. These DSSs are based on intuitive models of human decision making so that the computations of the system should be understandable to the user. Thus, intuitive DSSs serve as a guide through complex and computationally demanding decisions.

Currently, it appears that intuitive DSSs are more popular. Examples, such as the TADMUS DSS, are indicative of this approach and the desire of operators to retain a high degree of control over the decision making process. ESs may be sophisticated and generate high quality COAs but users will lack a clear basis for evaluating the output of an ES because their operation does not emulate human decision making. Intuitive DSSs also make better use of the computational power of computers than interface DSSs. They help users by performing additional decision making steps to reduce the workload of the human decision maker.

Attention should also be paid to emerging technologies that could expand the power of DSSs. VR technology, for example, offers the possibility of dramatically improving communications within the OR and between OR personnel on different platforms. This technology could allow

the development of VCCs to support unified tactical displays that multiple individuals could inspect and interact with. VCCs offer the possibility of simulated face-to-face communications, reference to a common display, and enhanced 3D displays.

9.3.3 SA

More research needs to be done to test current theories of SA and develop a better understanding of the processes of building SA. Some work has documented common errors of SA and factors that can limit a person's ability to achieve adequate SA. These results point to a number of important considerations for C2 and design of decision support. In particular, limits of attention and WM pose a serious problem in naval tactical decision making because of the vast amount of information a decision maker is confronted with at any given moment. So far, however, little research has addressed ways to control these factors.

9.3.4 HCI

One issue of cognitive fit and interface design concerns whether to present information in a graphic or text form because these support different cognitive strategies. Graphics are themselves analog and spatial and so support relational processing. Text consists of discrete information and is a good format to convey precise bits of data. Thus, a crucial aspect of the design process is determining what kinds of information the user will employ and how that information can best be conveyed.

Another technique to enhance cognitive fit is to create an embedded user model in the system interface. This is a cognitive and behavioural model of the user, specifying the kinds of cognitive strategies he or she will likely use, information requirements, and so on. The model controls interface features so that the system can anticipate the needs of the user and alter display format, command structure, and input protocols in advance of the user's needs. This technique has the promise to make interfaces even more responsive to user needs.

9.3.5 Expertise and Training

Although a great deal of evidence supports the stage-like development of expertise in many domains, research is needed to verify this theory for naval decision making. Aspects of C2, notably the levels of uncertainty and time pressure, could potentially make it a domain unlike chess or physics or other domains in which expertise has been examined.

Nevertheless, research demonstrating stages of expertise have implications for training, notably that training should provide extensive practice with realistic problems in realistic settings. This practice serves the joint function of providing the extensive knowledge base needed to develop expertise and helping trainees organize that knowledge around important concepts and problem types. Furthermore, practice is crucial for the transition from a declarative understanding of the domain to a procedural understanding that allows fast and relatively automatic responding to problems.

A number of training programs have already been developed to provide operationally relevant practice and instruction. CTT techniques involve training in the context of realistic scenarios with review and feedback designed to help trainees discover critical concepts and errors in their current strategies. These techniques make use of experiential learning, which has been found to promote better learning in complex, procedural domains than standard lecture

techniques on their own. Traditional techniques have a place in training programs, especially early on when complete novices rely on a declarative understanding of the new domain. As trainees progress, however, there is a growing mismatch between their increasingly procedural understanding and the content of lectures. Thus, experiential learning, supported by multimedia, becomes more effective as trainees gain experience.

9.4 Methodology

There is a vast array of methodologies available for the study of C2 and decision support. Research in these areas has drawn on the behavioural sciences, computer design, and HF to derive procedures and measures for assessing the effectiveness of C2 systems. Thus, there seems to be no well-defined set of "correct" or "appropriate" methodologies that should be employed in all cases for all purposes. Rather, researchers should continue to adopt a "pick and choose" approach to select methodologies that suit the interests of the moment and can illuminate the issues under consideration.

One danger to consider, however, is that researchers have tended to adopt a focus on C2 systems rather than taking a somewhat broader approach to consider the interactions of the human decision maker, the system, and the organization as an entity for study. Most techniques discussed in the literature assess the effectiveness of equipment or the effectiveness of a person using some equipment. This is not, in itself, a wrong approach because the system and the user are important components of C2. Understanding how each operates is essential. Taking a strictly reductionist approach, however, may not give a complete understanding of how well an organization, such as the crew of a CPF, will function. Interactions between individuals, between individuals and computer systems, and between teams and the organizational context can produce effects that will not be evident if one measures the performance of a single component.

Unfortunately, systemic measures that capture these kinds of interactions are rare and, presumably, difficult to derive. Further research should consider how to integrate all aspects of C2 (human, machine, and organization) to characterize the functioning at a ship level.

9.4.1 Design Process

Perhaps the best opportunity to develop a systemic methodology is to create a comprehensive design methodology that specifies how an equipment system will be developed from the conceptual to implementation stage. As such, it will specify the kinds of issues addressed.

A key design concept is that of user-centred design. In this methodology, the focus of design is to involve the user early and to continually obtain user input to define system requirements and refine design ideas. Thus, no stage of development should proceed without a plan to survey a representative group of future users and integrate their feedback with ideas of designers.

Related to user-centred design is the practice of rapid prototyping. This is the method of building representations of the system to some level that conveys the system's functionality and interface. The prototype serves as a concrete tool to solicit user input to help refine design concepts. To be useful, the prototyping effort must be iterative with a number of redesigns of the system based on user input.

A number of specific design methods have been proposed (e.g., Adelman, 1992; Cochrane & Foley, 1991; Schuffel, 1994; Webb et al., 1993). Selecting one of these is not critical but

clearly specifying what design methodology will be used before the system development project is initiated is crucial.

9.4.2 Research Tools

Simulation is one of the chief research tools used in the study of tactical decision making. Most research focuses on presenting participants with a realistic situation and set of tasks to capture the experience of operationally relevant tactical decision making. Two major kinds of simulation were identified in this literature review. The first, scenario-based, is by far the most commonly used. This technique involves scripting an event or series of events from the perspective of the participant. Thus, the participant is presented with sensor data corresponding to the data that would be available in the real situation. The data depicts the actions of hostile forces and other factors that are dictated by the scenario. The second technique is the cellular automata approach, which focuses on creating a simulation of the tactical domain in which the participant can interact with friendly and enemy forces in a realistic manner. This approach does not provide the researcher with complete control over all events in the simulation but does allow for rapid and inexpensive simulations of field exercises.

Neither approach by itself will serve every research need. Scenario-bases simulations provide a good level of control but are time consuming and their operational relevance depends highly on the researcher who developed them. Cellular automata simulations are quick and easy to conduct and allow for extensive practice in solving tactical problems but require a great deal of effort to initially produce. Furthermore, they typically present participants with data and interfaces that have little fidelity to operational conditions. Both approaches should be considered, along with a number of other techniques, including laboratory experiments, field exercises, and testing in operational settings.

Simulation techniques must be combined with methods for assessing performance. Subjective methods are common, in large part because of their simplicity. Subjective techniques, however, may be problematic because they depend on the ability of participants or raters to assess behaviours. Empirical techniques tend to be more objective but one must ensure validity and reliability.

9.4.3 Measures

A fair amount of effort has been devoted to developing measures that allow researchers to accurately assess the performance and value of C2 systems. This is not an easy task because some basic issues, such as the nature of effectiveness and the objectives of the system can be difficult to define in operational terms. That is why it is important to examine measurements prior to any evaluation of a system. Development of measurements, in fact, is probably best done as part of the system design process, in conjunction with requirements definition.

One issue that needs to be further addressed is whether to assess outcomes versus performance. Outcome measures can be easier to obtain and refer directly to the issues of interest – did the friendly force destroy the enemy, was the hostile aircraft shot down, etc. A problem with outcome measures is that the objectives of a mission are often not specified in sufficient detail to form a criterion for evaluating an outcome. A friendly force that captures a target can judged to have a positive or negative outcome based on factors such as levels of acceptable loses, expenditures of resources, time, and so on. Perhaps even more troublesome, outcomes

can be affected by numerous factors outside the system under consideration so that a positive outcome may not be informative about the system at all. For this reason, many researchers employ performance measures to assess the operation of the system. The value of performance measures, however, depends on the degree to which the researcher can specify the relevant aspects of performance and link them to operational objectives.

There is a range of measurement types (MOEs, MOFEs, MOPs, and physical parameters) that can be used. Which of these are appropriate depends on the goals of the assessment and the issues discussed above. Similarly, there are a number of techniques for gathering measures, including subjective, observational, and objective techniques. Further research should examine in more detail the relative merits of these approaches.

Given the changing role of the Navy, another area in need of further research is the development of MOEs for MOOTW. The Navy will be called upon to serve in non-traditional missions and unusual circumstances. There currently exist few, if any, measures to assess performance in these situations. Such measures, however, will be needed to determine how well C2 and DSSs perform in roles outside traditional warfare.

9.4.4 Analysis Techniques

One methodology that will be critical to any system development is CTA. Technically, a family of techniques, CTA allows researchers to describe, in detail, the cognitive and behavioural steps involved in performing complex tasks. User-centred design, requirements specification, and measurement of effectiveness all depend on an understanding of what the human decision maker is thinking and doing. CTA is, perhaps, one of the best ways to develop that understanding (although CWA also seems like a promising technique).

Numerous existing CTA techniques have been developed. Klein's CIM, for example, is a popular and well-developed example. The specific technique chosen, however, seems less important than the general principles. In fact, researchers should freely adapt CTA methods to suit the research issues and domain, treating CTA as a set of guidelines for eliciting participants' cognitive strategies rather than a hard-and-fast methodology.

There are a number of weaknesses, some practical, some conceptual, to CTA that need to be addressed. CTA is cost and labour intensive, making it difficult to employ. It generates vast amounts of data, which can actually hamper researchers who find themselves unable to thoroughly examine all of the data. CTA is ultimately a subjective technique, relying on participants to identify critical components of their own thinking. Researchers can deal with this problem to some extent by creating very specific queries to elicit theoretically relevant responses. Nevertheless, researchers must ultimately trust that participants have accurate access to their cognitive processes, which is almost certainly never completely the case.

9.4.5 Team Research

Development of methodologies to assess team performance and teamwork probably lags behind the development of theories of teamwork. Certainly, this area presents additional challenges because researchers must deal not only with the performance of all the individual team members but their interactions as well. A number of methods to assess team performance have been developed, which generally focus on coordination and communication aspects of teamwork and seek to quantify how well team members are able to pass

information and coordinate responses. These methods may be hampered by a lack of any agreed upon means to analyze communications.

Measuring shared mental models is another key aspect of assessing teamwork. Theories of teamwork point to this as a major determinant of how well team members will be able to work together. Techniques for measuring shared mental models, however, are very primitive and have not solved the theoretical issue of determining the overlap between individuals' mental models. This will be a difficult problem to solve because mental models consist of more than just the information they contain. The organization and conceptual structures of mental models are also key. What is needed is an analytic technique to quantify the extent to which individuals possess the same conceptual structure as well as the same information about a task.

9.5 Guidelines

Relatively few practical guidelines were obtained for this literature review. This probably reflects, to some extent, the focus of the literature search. Although the term guideline was included a search term to obtain articles for review, most search words addressed theoretical and empirical issues. A more focused review of the literature could likely identify more relevant guidelines. Nevertheless, this may be a less developed area than theoretical and empirical studies.

9.6 Conclusions

This literature review has surveyed basic and applied research related to single ship C2 to identify issues and problems for the implementation of decision support in the upgrade of the HALIFAX class. From this, it is clear that a great deal of work remains to be done to fully characterize the decision support needs for the HALIFAX class. This review has identified a number of major areas, including decision making, SA, teamwork, training, and HCI, as critical to C2 and decision support. It is clear from the literature that each of these areas is complex enough that future research is needed to verify and adapt theories and guidelines to suit the specific tasks and conditions of the HALIFAX class.

From the review, we have identified a number of major conclusions.

1. Future literature reviews should consider relevant topics in the broader, *non-naval literature* on Situation Awareness (SA) and decision making, taking care to avoid extensive, unproductive literature searches by focusing on only highly related research.
2. The future naval environment will place greater demands on naval decision makers who will be required to act rapidly, under high risk, with uncertain information, in novel and unfamiliar situations.
3. Tactical decision makers primarily use intuitive decision making strategies (e.g., recognition and story generation), indicating a need for decision support to overcome the weaknesses of these strategies (i.e., no analysis of the quality of solutions, dependence on familiarity of situations).
4. Research has tended to ignore mission planning and other preparation activities that are part of C2 but these areas are crucial and may be supported by analytic theories of decision making.
5. Further research should be conducted to clarify the processes by which people develop SA, the specific requirements for SA support, and the relation of SA to decision making.
6. Naval C2 places a great emphasis on teamwork (within and between ships), indicating a need for further study of how teams coordinate, communicate, allocate tasks and resources, achieve team SA, and make decisions.

7. Researchers should follow user-centred design methods, involving users from the start to ensure that system requirements and design concepts will help users achieve their goals.
8. Researchers should make use of rapid prototyping methods to increase the effectiveness of user-centred design and provide users with concrete models of system features.
9. There is no single best method to evaluate system or human performance and researchers should adopt a broad research approach that makes use of experimental, observational, and simulation studies, conducted with high and low levels of fidelity to operational conditions (a demanded by empirical goals), and employing a range of measures of effectiveness.
10. Researchers should explore emerging technologies and techniques that can potentially enhan the Navy's ability to evaluate new C2 systems.
11. There is a need to develop team research techniques and methods to support research in teamwork and to integrate theories of teamwork with theories of decision making.
12. Survey existing Decision Support Systems (DSSs) to identify concepts, techniques, and HCI features to determine which are directly applicable to the HALIFAX class.
13. Survey scientific, engineering, and computer science literatures to identify practical guidelin for DSSs, C2 systems, and HCI.

Overall, there are no simple answers to the questions of how OR personnel understand situations and make decisions, what factors affect their abilities, and what are their decision support requirements. The complexities of the C2 domain and human cognition make it impossible for any current theories to completely explain human performance completely or in detail. This underscores the need to adopt a user-centred approach and constantly examine and refine theoretical and design concepts with respect to the specific tasks and personnel of naval C2 aboard the CPF.

10. References

1998). GUI Design Methodology. http://www.dordt.edu:457/VTCLG/vtclgN.style_method.html

(1998). HCI Guidelines. http://www.cic13.lanl.gov/guiguide/Guiguid4.html

(1998). IBM Human-Computer Interaction: UI Architecture and Design. http://www.ibm.com/imb/hci/designer/uiarchnt.html

(1998). Industry Guidelines. http://www.dordt.edu:457/VTCLG/vtclgD.style_standards.html

(1998). What Makes a Good User Interface? http://www.dordt.edu:457/VTCLG/vtclgN.style_goodui.html

Adelman, L. (1992). *Evaluating Decision Support and Expert Systems*. New York: John Wiley & Sons.

Adelman, L., Cohen, M.S., Bresnick, T.A., Chinnis, J.O., Laskey, K.B. (1993). Real-Time Expert System Interfaces, Cognitive Processes, and Task Performance: An Empirical Assessment. *Human Factors*, 35(2): 243-261.

Anderson, J. R. (1995). *Cognitive Psychology and its Implications (4th Ed.)*. New York: W. H. Freeman and Company.

Anderson, J. R., & Fincham, J. M. (1994). Acquisition of Procedural Skills From Examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20: 1322-1340.

Anderson, J.R., Fincham, J.M., Douglass, S. (1997). The Role of Examples and Rules in the Acquisition of a Cognitive Skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23(4):932-945.

Anderson, S.E. (1990). *Functional Specification for a Generic C3I Workstation*. Naval Postgraduate School, Monterey, CA: 271pp. Master's Thesis.

Anzai, Y. (1984). Cognitive Control of Real-time Event-driven Systems. *Cognitive Science*, 8(3): 221-254.

Baddeley, A. (1990). *Human Memory: Theory and Practice*. Boston, MA: Allyn and Bacon.

Baker, D.P., Salas, E. (1992). Principles for Measuring Teamwork Skills. *Human Factors:* 34(4): 469-475.

Baker, D.P., Salas, E. (1997). Principles for Measuring Teamwork: A Summary and Look Toward the Future. In M.T. Brannick, E. Salas, C. Prince (Eds.) *Team Performance Assessment and Measurement: Theory, Methods, and Applications*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ: 331-355.

Beach, L. R. (1990). *Image Theory: Decision Making in Personal and Organizational Contexts*. West Sussex, England: Wiley.

Best, J. B. (1995). *Cognitive Psychology (4th Ed.)*. Minneapolis, MN: West Publishing.

Bostow, D. E., Kritch, K. M., & Tompkins, B. F. (1995). Computers and Pedagogy: Replacing Telling with Interactive Computer-Programmed Instruction. *Behavior Research Methods, Instruments & Computers*, 27: 297-300.

Bowers, C.A., Braun, C.C., Morgan Jr., B.B. (1997). Team Workload: Its Meaning and Measurement. In M.T. Brannick, E. Salas, C. Prince (Eds.) *Team Performance Assessment and Measurement: Theory, Methods, and Applications.* Lawrence Erlbaum Associates, Mahwah, NJ: 85-108.

Brannick, M.T., Prince, A., Prince, C., Salas, E. (1995). The Measurement of Team Process. *Human Factors* 37(3): 641-651.

Brecke, F. H., & Garcia, S. K. (1995). Training Methodology for Logistic Decision Making. *USAF AMRL Technical Report*, AH/HR-TR-1995-0098.

Bryant, D. J., Webb, R. D. G., & McLean, D. N. (1998). *Literature Survey for Issues in Naval Decision Support: Phase I.* Report for Defence and Civil Institute of Environmental Medicine.

Campion, J., Brander, G., Koritsas, E. (1996). *Principles of Decision Aiding: Aid Design Concepts and Laboratory Demonstrator Development.* Defence Evaluation Research Agency, Farnborough, Hants

Cannon-Bowers, J. A., & Bell, H. H. (1997). Training Decision Makers for Complex Environments: Implications of the Naturalistic Decision Making Perspective. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ: 99-110.

Cannon-Bowers, J.A., Oser, R., Flanagan, D.L. (1992). Work Teams in Industry: A Selected Review and Proposed Framework. In R.W. Swezey and E. Salas (Eds) *Teams: Their Training and Performance.* Ablex Publishing, Norwood, NJ: 355-377.

Cannon-Bowers, J.A., Salas, E. (1997). A Framework for Developing Team Performance Measures in Training. In M.T. Brannick, E. Salas, C. Prince (Eds.) *Team Performance Assessment and Measurement: Theory, Methods, and Applications.* Lawrence Erlbaum Associates, Inc., Mahwah, NJ: 45-62.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E., Volpe, C.E. (1995). Defining Competencies and Establishing Team Training Requirements. In R. Guzzo and E. Salas (Eds). *Team Effectiveness and Decision Making in Organizations*. Jossy Bass, San Francisco, CA: 333-380.

Casey, C. (1996). Incorporating Cognitive Apprenticeship in Multi-Media. *Educational Technology Research & Development*, 44: 71-84.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic Reasoning Schemas. *Cognitive Psychology,* 17: 391-416.

Cochran, P. L. Foley, M. J. (1991). *Assessment of the Marine Tactical Command and Control System (MTACCS).* Naval Postgraduate School, Monterey, CA.

Cohen, M.S., Freeman, J.T., Thompson, B.B. (1997). Integrated Critical Thinking Training and Decision Support for Tactical Anti-Air Warfare. *Proceedings of Third International Symposium on Command and Control Research and Technology*. National Defense University, Washington, DC: 365-377.

Colton, R. A., & Gange, R. H. (1993). *Including the Human Element in Design of Command and Control Decision Support Systems - The KOALAS Concept*, Naval Research Lab., Washington, DC.

Cyrus, A.W. (1991). *Measuring the Effectiveness of Information Systems*. Naval Postgraduate School, Monterey, CA: Master's Thesis.

Damper, R. I., & Wood, S. D. (1995). Speech Versus Keying in Command and Control Applications. *International Journal of Human-Computer Studies*, 42(3): 289-305.

De Groot, A. D. (1965). *Thought and Choice in Chess*. The Hague, Netherlands: Mouton.

Dockery, J.T., Hill, J.M. (1997). Virtual Command Center. *Proceedings of Third International Symposium on Command and Control Research and Technology*. National Defense University, Washington, UK: 61-68.

Dordick, M. (1997). A Computerized Simulation for ESL Composition Students: Building a House. *Simulation & Gaming*, 28: 98-106.

Drillings, M., & Serfaty, D. (1997). Naturalistic Decision Making in Command and Control. In G. Klein and C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 71-80.

Driskell, J. E., & Salas, E. (1992). Collective Behavior and Team Performance. *Human Factors* 34(3): 277-288.

Driskell, J.E., Salas, E. (1992). Can You Study Real Teams in Contrived Settings? The Value of Small Group Research to Understanding Teams. In R.W. Swezey and E. Salas (Eds) *Teams: Their Training and Performance*. Ablex Publishing, Norwood, NJ: 101-124.

Elliott, L.R., Schiflett, S.G., Dalrymple, M.A. (1996). Indices of Performance in AWACS Teams: Constructs and Measures. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 686-693.

Endsley, M. (1997). The Role of Situation Awareness in Naturalistic Decision Making. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 269-283.

Endsley, M.R. (1995a). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors* 37(1): 32-64.

Endsley, M.R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors* 37(1): 65-84.

Fallesen, J. J., & Pounds, J. (1998). Identifying and Testing a Naturalistic Approach for Cognitive Skill Training. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Federico, P.-A. (1995). Expert and Novice Recognition of Similar Situations. *Human Factors*, 37 (1): 105-122.

Federico, P.-A. U., R. R.; Van de Wetering, B. L.; Tomlinson, C. I.; Long, D. J. (1991). *Human-Computer Interfaces for Tactical Decision Making, Analysis, and Assessment Using Artificially Intelligent Platforms. Volume 1. Software Design and Database Descriptions for BATMAN and ROBIN.*, Navy Personnel Research and Development Center, San Diego, CA.

Fleishman, E.A., Zaccaro, S.J. (1992). Towards a Taxonomy of Team Performance Functions. In R.W. Swezey and E. Salas (Eds). *Teams: Their Training and Performance*. Ablex Publishing, Norwood, NJ: 31-56.

Flin, R. (1998). Applying NDM to Critical Incident Management. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Foster, T.A. (1992). *C2 Information Management: Data Fusion and Track ID's in a Multiple Sensor Environment*. Naval Postgraduate School, Monterey, CA: 62pp. Master's Thesis.

Fowlkes, J.E., Lane, N.E., Salas, E., Franz, T., Oser, R. (1994). Improving the Measurement of Team Performance: The TARGETs Methdology. *Military Psychology*. 6(1): 47-61.

Freeman, J.T., Cohen, M.S. (1996). Training for Complex Decision-Making: A Test of Instruction Based on the Recognition/Metacognition Model. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 260-271.

Gabrielson, D. (1995). *Battle Group Stationing Algebraic Modeling System: An Anti-Air Warfare Tactical Decision Aid Methodology*. Naval Postgraduate School, Monterey, CA.

Galson, N. J., & Oliker, L. R. (1976). Assessment of Experiential Learning in Business. *Alternative Higher Education,* 1: 33-42.

Gardner, S.B., Balash, J., Callihan, H., Saverino, M. (1996). A Virtual Communications Environment (VCE) for Command and Control of Complex Systems. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 563-570.

Garner, K.T. & Assenmacher, T.J. (1997). *Situational Awareness Guidelines.*, Naval Air Warfare Center Aircraft Div., Patuxent River, MD.; Naval Air Systems Command, Washington, DC.

Geiselman, R.E., Landee-Thompson, B., Samet, M.G. (1986). A Selective Callup System for Managing Tactical Information on Graphic Displays. Special Issue: Information Technology for Command and Control. *IEEE Transactions on Systems, Man & Cybernetics*, 16(6): 901-908.

Gick, M. L., & Holyoak, K. J. (1983). Schema Induction Analogical Transfer. *Cognitive Psychology,* 15: 1-38.

Gigerenzer, G., & Hug, K. (1992). Domain-Specific Reasoning: Social Contracts, Cheating, and Perspective Change. *Cognition,* 43: 127-171.

Gilhooly, K. J. (1989). Human and Machine Problem Solving: Toward a Comparative Cognitive Science. In K. J. Gilhooly (Ed.) *Human and Machine Problem Solving*. Plenum Press, New York: 1-12.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M.T.H. Chi, R. Glaser, & M. Farr (Eds.), *The Nature of Expertise*. Hillsdale, NJ: Erlbaum.

Glover, S.M., Prawitt, D.F., Spilker, B.C. (1997). The Influence of Decision Aids on User Behavior: Implications for Knowledge Acquisition and Inappropriate Reliance. *Organizational Behavior and Human Decision Processes*. 72(2):232-255.

Gugurty, L. J. (1997). Situation Awareness During Driving: Explicit and Implicit Knowledge in Dynamic Spatial Memory. *Journal of Experimental Psychology: Applied*, 3: 42-66.

Hair, D. C. & Pickslay, K. (1992). *Development of Computer Support for Naturalistic Decision-Making.*, Performing Org.: Naval Command, Control and Ocean Surveillance Center, San Diego, CA. RDT and E Div. (Code: DODXA).

Hair, D. C. & Pickslay, K. (1993). *User Interface Issues in Real Time Decision Support Systems.*, Naval Command, Control and Ocean Surveillance Center, San Diego, CA. RDT and E Div.

Hakala, J.W. (1995). *Combat Simulation Modeling in Naval Special Warfare Mission Planning.* Naval Postgraduate School, Monterey, CA: 89pp. Master's Thesis.

Halpin, S. M. (1992). *Army Command and Control Evaluation System (ACCES): A brief description.* US Army Research Institute for the Behavioral and Social Sciences. Alexandria, VA.

Hart, J.H. (1988). *Analysis and Design of a Microcomputer-Based Decision Support System for the US Navy's Shipboard Tactical Action Officer.* Naval Postgraduate School, Monterey, CA: 84pp. Master's Thesis.

Hays, W. L. (1988). *Statistics (4th Ed.).* New York, NY: Holt, Rinehart andWinston.

Herz, B., & Merz, W. (1998). Experiential Learning and the Effectiveness of Economic Simulation Games. *Simulation & Gaming*, 29: 238-250.

Heward, W. L. (1997). Four Validated Instructional Strategies. *Behavior and Social Issues*, 7: 43-51.

Hill, M. (1996). Scenario Assisted Requirements Identification. *Proceedings of Second International Symposium on Command and Control Research and Technology.* Market Bosworth, Warwickshire, UK: 281-287.

Hoffman, R. R., Crandeell, B., Shadbolt, N. (1998). Use of the Critical Decision Method to Elicit Expert Knowledge: A Case Study in the Methodology of Cognitive Task Analysis. *Human Factors*, 40: 254-276.

Hollenbeck, J. R., Ilgen, D.R., et al. (1995). Team Performance on Monitoring Tasks: An Examination of Decision Errors in Contexts Requiring Sustained Attention. *Journal of Applied Psychology* 80(6): 685-696.

Hollenbeck, J.R., Ilgen, D.R., Weissbein, D. (1996). Information Management and Decision Making Following Unusual Events: Analysis at the Team and Individual Level. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 306-321.

Hollingshead, A.B., McGrath, J.E. (1995). Computer-Assisted Groups: A Critical Review of the Empirical Research. In R. Guzzo and E. Salas (Eds). *Team Effectiveness and Decision Making in Organizations.* Jossy Bass, San Francisco, CA: 46-78

Howell, W.C., Lane, D.M., Holden, K.L. (1988). *Human Cognition and Information Display in C3I System Tasks.* Report prepared for: US Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA: 42pp.

Howell, W.C., Lane, D.M., Laux, L.F., Anderson, L.A., Holden, K.L. (1993). *Visual Display Principles for C3I System Tasks.* Report prepared for: US Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA: 48pp. ARI Research Note 93-19.

Hughes, Jr., W.P. (1996). Littoral Operations, Maneuver Warfare, and Naval Doctrine. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 25-28.

Hutchins, S. G. (1997). Decision-Making Errors Demonstrated by Experienced Naval Officers in a Littoral Environment. [References]. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 207-215.

Hutchins, S. G., & Kowalski, J. T. (1993). *Tactical Decision Making Under Stress - Preliminary Results and Lessons Learned.*, Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

Hutchins, S.G. (1996). *Principles for Intelligent Decision Aiding*. Naval Command, Control and Ocean Surveillance Center, RDT&E Division, San Diego, CA: 33pp. Technical Report 1718.

Hutchins, S.G., Duffy, L.T. (1993). *Decision Making Evaluation Facility for Tactical Teams*. Naval Command, Control and Ocean Surveillance Center (NCCOSC), San Diego, CA. Report No. AD-A265 058.

Hutchins, S.G., Morrison, J.G., Kelly, R.T. (1996). Principles for Aiding Complex Military Decision Making. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 186-203.

Ilgen, D.R., Major, D.A., Hollenbeck, J.R., Sego, D.J. (1995). Raising an Individual Decision-Making Model to the Team Level: A New Research Model and Paradigm. In R. Guzzo and E. Salas (Eds). *Team Effectiveness and Decision Making in Organizations*. Jossy Bass, San Francisco, CA: 113-148.

Jarvenpaa, S. L. (1989). The Effect of Task Demands and Graphical Format on Information-Processing Strategies. *Management Science* 35(3): 285-303.

Jarvenpaa, S. L., & Dickson, G. W. (1988). Graphics and Managerial Decision Making: Research Based Guidelines. *Comm.* ACM, 31: 764-774.

Johnston, J.H., & Cannon-Bowers, J.A. (1996). A Conceptual Framework and Recommendations for Integrating Training with a Decision Support System. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 233-240.

Jones, D.G., Endsley, M.R. (1996). Sources of Situation Awareness Errors in Aviation. *Aviation, Space and Environmental Medicine* 67(6): 507-512.

Kaempf, G. L., & Orasanu, J. (1997). Current and Future Applications of Naturalistic Decision Making in Aviation. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making.* Lawrence Erlbaum Associates Inc, Mahwah, NJ: 81-90.

Kahneman, D., & Tversky, A. (1972). Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology,* 3: 430-454.

Kalus, A., Ashford, A.J., Cook, C.A., Sheppard, C. (1996). An Introduction to Soar. *Proceedings of Second International Symposium on Command and Control Research and Technology.* Market Bosworth, Warwickshire, UK: 288-297.

Kelly, R.T., Hutchins, S.G., Morrison, J.G. (1996). Decision Processes and Team Communications with a Decision Support System. *Proceedings of Second International Symposium on Command and Control Research and Technology*. Market Bosworth, Warwickshire, UK: 562-567.

Kirschenbaum, S.S. (1992). Influence of Experience on Information-Gathering Strategies. Journal of *Applied Psychology*. 77(3): 343-352.

Klein, G. (1992). *Decision Making in Complex Military Environments*. Klein Associates, Fairborn, OH: Final Contract Summary Report, prepared for Naval Command, Control and Ocean Surveillance Center, San Diego, CA.

Klein, G. (1997). An Overview of Naturalistic Decision Making Applications. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 49-59.

Klein, G., G. L. Kaempf, et al. (1997). Applying Decision Requirements to User-Centered Design. *International Journal of Human-Computer Studies* 46(1): 1-15.

Klein, G., Wolf, S., Militello, L., Zsambok, C. (1995). Characteristics of Skilled Option Generation in Chess. *Organizational Behavior and Human Decision Processes*. 62(1): 63-69.

Kozlowski, S. W. J., Toney, R. J., Weissbein, D. A., Mullins, M. E., & Brown, K. G. (1998). Developing Adaptive Expertise. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Kozlowski, S.W.J., Gully, S.M. (1996). TEAMS/TANDEM: Examining Skill Acquisition, Adaptability, and Effectiveness. In J. Vancouver and A. Williams (Chairs), *Using Computer Simulations to Study Complex Organizational Behavior*. Symposium conducted at the Annual Convention of the Academy of Management Association, Cincinnati, OH: 33pp.

Kraiger, K., Wenzel, L.H. (1997). Conceptual Development and Empirical Evaluation of Measures of Shared Mental Models as Indicators of Team Effectiveness. In M.T. Brannick, E. Salas, C. Prince (Eds.) *Team Performance Assessment and Measurement: Theory, Methods, and Applications*. Lawrence Erlbaum Associates Inc., Mahwah, NJ: 63-84.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional Scaling*. Beverly Hills, CA: Sage.

Kulik, J. A., & Kulik, C. C. (1988). Timing of Feedback and Verbal Learning. *Review of Educational Research,* 58: 79-97.

Landauer, T. K., & Bjork, R. A. (1978). Optimum Rehearsal Patterns and Name Learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*, London: Academic Press: 625-632.

Lee, C. S., Rutecki, G. W., Whittier, F. C., Clarrett, M. R., & Jarjoura, D. (1997). A Comparison of Interactive Computerized Medical Education Software with a More Traditional Teaching Format. *Teaching and Learning in Medicine*, 9: 111-115.

Lee, M. W. (1984). Strategies for Teaching Young Children: Guides for Improving Instruction. *Early Child Development and Care*, 14: 189-200.

Leedom, D. K., Adelman, L., & Murphy, J. (1998). Critical Indicators in Naturalistic Decision Making. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Leroy, D., Hofmman, D. (1996). Criterion of C3I's Effectiveness. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 396-403.

Lesgold, A.M., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a Complex Skill: Diagnosing X-ray Pictures. In M.T.H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Libby, T.E. (1997). The Navy's C4ISR Architecture Process. *Proceedings of Third International Symposium on Command and Control Research and Technology*. National Defense University, Washington, DC: 415-425.

Lipshitz, R., Shaul, O.B. (1997). Schemata and Mental Models in Recognition-Primed Decision Making. [References]. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ:293-303.

Liu, L., Schneider, P., & Miyazaki, M. (1997). The Effectiveness of Using Simulated Patients Versus Videotapes of Simulated Patients to Teach Clinical Skills to Occupational and Physical Therapy Students. *Occupational Therapy Journal of Research*, 17: 159-172.

Loftus, EE. F. (1980). *Memory*. Reading, MA: Addison-Wesley.

Lossau, K., Haug, D. (1996). A Force Independent Object Representation to Support Course of Action Planning and Visualization. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 527-535.

Manning, S N. (1991). *Conceptual Design for the Telecommunications Emergency Decision Support System (TEDSS).*, Naval Postgraduate School, Monterey, CA: 89pp. Master's Thesis.

Mason, D. E. (1995). *Identifying Measures of Effectiveness for Marine Corps C4I Systems.*, Naval Postgraduate School, Monterey, CA.

Mathieson, G., Grenham, W., Miller, Maj G. (1996). Making the Most of Scenarios Through Study-Specific Characterisation. *Proceedings of Second International Symposium on Command and Control Research and Technology*. Market Bosworth, Warwickshire, UK: 268-280.

Matthews, M. L., Webb, R. D. G., & McCann, C. (1997). *A Framework for Evaluation of Military Command and Control Systems*. Report to Defence and Civil Institute of Environmental Medicine.

Mayer, R. E. (1989). Human Nonadversary Problem Solving. In K. J. Gilhooly (Ed.) *Human and Machine Problem Solving*. Plenum Press, New York:1-12.

McCall, R. B. (1980). *Fundamental Statistics for Psychology (3rd Ed.)*. New York, NY: Harcourt Brace Jovanovich.

McMenamin, LtCol. J.J. (1992). *Operational Decision Making: The Impact of Time and Information*. Naval War College, Newport, RI: 16pp.

Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided Learning by Doing. *Cognition and Instruction,* 13: 315-372.

Militello, L. G. (1998). Capturing Expertise. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Miller, A.E., Blanks, K.S., Le, H. (1997). Command, Control, Communications, Computers and Intelligence (C4I) Functional Assessment Methodology - Real World COTS Instrumentation, Quantitative, and Web-based Assessment Techniques. *Proceedings of Third International Symposium on Command and Control Research and Technology*. National Defense University, Washington, DC:497-507.

Miller, T. E., Woods, D. D. (1997). Key Issues for Naturalistic Decision Making Researchers in System Design. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 141-149.

Miller, T.E., Wolf, S.P., Thorsden, M.L., Klein, G. (1992). *A Decision-Centered Approach to Storyboarding Anti-Air Warfare Interfaces*. Report prepared for: Naval Command Control and Ocean Surveillance Center, San Deigo, CA: 103pp.

Miserandino, M. (1998). Those Who Can Do: Implementing Active Learning. *APS Observer,* 11: 24-26.

Morrison, J. G., Kelly, R. T., Moore, R. A., & Hutchins, S. G. (in press). Implications of Decision Making Research for Decision Support and Displays. To appear in J. A. Cannon-Bowers & E. Salas (Eds.), *Decision Making Under Stress: Implications for Training and Simulation.*

Najjar, L. J. (1998). Principles of Educational Multimedia User Interface Design. *Human Factors,* 40: 311-323.

Newell, A. (1990). *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Newell, A., & Simon, H. A. (1961). GPS: A Program That Simulates Human Thought. In H. Billing (Ed.), *Lernende Automaten,* Munich: R. Oldenbourg: 109-124.

Noakes, B., Winkles, J., Davis, R., Webb, T. (1996). Distributed Problem Solving and Decision Making in Military Command and Control. *Proceedings of Second International Symposium on Command and Control Research and Technology*. Market Bosworth, Warwickshire, UK: 312-325.

Norman, D.A. (1981). Categorization of Action Slips. *Psychological Review.* 88(1):1-15

O'Neill, J. (1996). The Role of Problem-Solving in C3I Systems. *Proceedings of Second International Symposium on Command and Control Research and Technology*. Market Bosworth, Warwickshire, UK: 348-355.

Omodei, M., McLennan, J., & Wearing, A. (1998). Integrating Field Observation and Computer Simulation to Analyze Command Behavior. Paper presented at *The Human in Command Conference,* Kingston: NATO RTA.

Pascual, R., Henderson, S. (1997). Evidence of Naturalistic Decision Making in Military Command and Control. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 217-226.

Pennington, N., & Hastie, R. (1986). Evidence Evaluation in Complex Decision Making. *Journal of Personality and Social Psychology,* 51: 242-258.

Pigeau, R., & McCann, C. (1995). Putting 'Command' back into Command and Control. Paper presented at the *Command and Control Conference,* Canadian Defence Preparedness Association, Ottawa, Ontario.

Pigeau, R., & McCann, C. (1998). Re-defining Command and Control. Paper presented at The *Human in Command Conference*, NATO RTO Workshop, Kingston, Ontario.

Pliske, R., McCloskey, M., & Klein, G. (1998). Facilitating Learning From Experience: An Innovative Approach to Decision Skills Training. *Fourth Conference on Naturalistic Decision Making*. Warrenton, VA: Klein Associates.

Radtke, P.H., Frey, P. (1996). PC-Based Part-Task Training for Team Decision Making: Recommendations and Guidelines. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 241-259.

Ramamurthy, K., Kind, W. R., & Premkemar, G. (1992). User Characteristics – DSS Effectiveness Linkage: An Empirical Assessment. *International Journal of Man-Machine Studies* 36: 469-505

Randel, J.M., Pugh, H.L., et al. (1996). Differences in Expert and Novice Situation Awareness in Naturalistic Decision Making. *International Journal of Human-Computer Studies* 45(5): 579-597

Randel, J.M., Pugh, H.L., Wyman, B.G. (1996). *Methods for Conducting Cognitive Task Analysis for a Decision Making Task.* Navy Personnel Research and Development Center, San Diego, CA: 30pp. Report NPRDC-TN-96-10.

Rapoport, A., & Wallsten, T. S. (1972). Individual Decision Making Behavior. *Annual Review of Psychology,* 23: 131-176.

Ray, W. J., Ravizza, R. (1985). *Methods Toward a Science of Behavior and Experience*. Belmont, CA: Wadsworth.

Resnick, H. (1994). Electronic Tools for Education and Training. *Computers in Human Services*, 11: 237-244.

Robson, S. (1997). Requirements Definition and Effectiveness Assessment. *Proceedings of Third International Symposium on Command and Control Research and Technology*. National Defense University, Washington, DC: 465-482.

Roth, E.M. (1997). Analysis of Decision Making in Nuclear Power Plant Emergencies: An Investigation of Aided Decision Making. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 175-182.

Rouse, W. B. (1980). *Human Detection and Diagnosis of System Failures*. New York: Plenum Press.

Rouse, W. B. (1985). Supervisory Control and Display Systems. In J. Zeidner (Ed.), *Human Productivity Enhancement*. New York: Praeger.

Rouse, W.B., Cannon-Bowers, J.A., Salas, E. (1992). The Role of Mental Models in Team Performance in Complex Systems. *IEEE Transactions on Systems, Man & Cybernetics,* 22(6): 1296-1308.

Rummel, B.K., (1995). *Subjective Evaluation of Human-Computer Interface Options for a Tactical Dislay Support System*. Naval Command, Control and Ocean Surveillance Center (NCCOSC), San Diego, CA. Report No. TR 1698.

Schuffel, H. (1994). *Human Factors in Bridge Operations - Decision-Support at Future Shipbridges (Interim Report).*, Human Factors Research Inst. TNO, Soesterberg (Netherlands).; Stichting Coordinatie Maritiem Onderzoek, Rotterdam (Netherlands).

Sengupta, K., Berigan, M., Kleinman, D.L., Serfaty, D. (1996). Architectures for Command and Control Organizations: Dimensions of Task and Organizational Structure. *1996 Command and Control Research and Technology Symposium,* Naval Postgraduate School, Monterey, CA: 281-287.

Serfaty, D., MacMillan, J., Entin, E.E., Entin, E.B. (1997). The Decision Making Expertise of Battle Commanders. In G. Klein & C. E. Zsambok (Eds.) *Naturalistic Decision Making.* Lawrence Erlbaum Associates Inc, Mahwah, NJ: 233-246.

Sovereign, M.G. (1996). Warfare in the Information Age: Vision 2010 and Changes in Joint C4 Doctrine. *1996 Command and Control Research and Technology Symposium*, Naval Postgraduate School, Monterey, CA: 141-147.

Speed, L. S. (1994). *Operational Commander and Staff - Procedures for Effective Decision Making.*, Naval War Coll., Newport, RI. Dept. of Operations.

Sternberg, R. J. (1996). *Cognitive Psychology*. Fort Worth, TX: Harcourt Brace College Publishers.

Stevens, W.K. Decker, W.L., Gagnon, C.M. (1996). Representation of Command and Control (C2) Information Warfare (IW) in Military Simulations. *Proceedings of Second International Symposium on Command and Control Research and Technology.* Market Bosworth, Warwickshire, UK: 240-255.

Su, D., Wei-Ming, Z., Wei-Dong, C. A Research on the Method of C3I System Evaluation. 497-500.

Sundin, Colonel C. (1996). ROLF – Mobile Joint Command and Control System for the Year 2010. A Hypothesis Under Development. *Proceedings of Second International Symposium on Command and Control Research and Technology.* Market Bosworth, Warwickshire, UK: 176-189.

Swezey, R.W., Salas, E. (1992). Guidelines for Use in Team-Training Development. In R.W. Swezey and E. Salas, *Teams: Their Training and Performance.* Ablex Publishing, Norwood, NJ: 219-245.

Tulving, E. (1983). *Elements of Episodic Memory*. London: Oxford University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology,* 5: 207-232.

Urban, C.D. (1990). *Design and Evaluation of a Tactical Decision Aid.* Naval Ocean Systems Center, San Diego, CA.

Vessey, I. (1991). Cognitive Fit: A Theory-Based Analysis of the Graphs versus Tables Literature. *Decision Sciences* 22: 219-240.

Vicente, K. J., & Rasmussen, J. (1992). Ecological Interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4): 589-606.

Vincent, M. (1976). Gaming and Simulation with Particular Reference to the Case-Study: How Well do you Use the Case Method? *Educational Studies*, 2: 111-120.

Vineberg, M., & Bryam, J. K. (1994). *Developing the Joint Maritime Command Information System Architecture Top Down*. Naval Command and Control and Ocean Surveillance Center.

Volpe, C.E., et al (1996). The Impact of Cross-Training on Team Functioning: An Empirical Investigation. *Human Factors,* 38(1): 87-100.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. New York: Wiley.

Wagner, D.H. (1989). *Naval Tactical Decision Aids*. Naval Postgraduate School. Report No. NPS55-89-011: 275pp.

Wall, A. N. (1997). *Situational Awareness in Military Operations Other Than War - A Look at Measures of Effectiveness in Humanitarian Assistance Operations*. Naval War College, Newport, RI.

Weaver, J.L., Bowers, C.A., Salas, E., Cannon-Bowers, J.A. (1995). Networked Simulations: New Paradigms for Team Performance Research. *Behavior Research Methods, Instruments, and Computers*, 27(1): 12-24.

Webb, R. D. G., Matthews, M. L., Greenley, M., & Burns, C. M. (1993). *Survey of C2 Evaluation Methods*. Department of National Defence, Canada.

Webb, R.D.G., & McLean, D. N. (1997). *Deficiencies in Command and Control Support to the Command Team in the Halifax Class*. Report for Defence and Civil Institute of Environmental Medicine.

Wiley, J. (1998). Expertise as Mental Set: The Effects of Domain Knowledge in Creative Problem Solving. *Memory and Cognition*, 26: 716-730.

Wohl, J., Serfaty, D., et al. (1988). Human Cognitive Performance in Antisubmarine Warfare: Situation Assessment and Data Fusion. *IEEE Transactions on Systems, Man & Cybernetics* 18(5): 777-786.

Wong, L.Y., et al (1989). *An Entity-Relationship Approach to Navy Command and Control Conceptual Data Modeling for Decision Support*. Naval Ocean Systems Center, San Diego, CA. *microfiche copy only.*

Xiao, Y., Milgram, P. (1998). Dare I Embark on a Field Study? Toward an Understanding of Field Studies. http://audio.ab.umd.edu/field/field.html

Xiao, Y., Milgram, P., Doyle, D.J. (1997). Capturing and Modelling Planning Expertise in Anesthesiology: Results of a Field Study. In G. Klein, C.E. Zsambok (Eds.) *Naturalistic Decision Making.*. Lawrence Erlbaum Associates Inc, Mahwah, NJ: 197-205.

Zachary, W.W., Ross, L. (1991). Enhancing Human-Computer Interaction Through Use of Embedded COGNET Models. *Proceedings of the Human Factors Society 35th Annual Meeting*. San Francisco, CA: 425-429.

Zechmeister, E. B., & Nyberg, S. E. (1982). *Human Memory: An Introduction to Research and Theory*. Monterey, CA: Brooks/Cole.

Ziems, D., & Neumann, G. (1997). Using Students' Knowledge to Generate Individual Feedback: Concept for an Intelligent Educational System on Logistics. *Journal of Artificial Intelligence in Education*, 8: 89-112.

ANNEX A:

To "Literature Survey for Issues in Naval Decision Support: Phase II"

Team Measures for AWACS crews (from Elliot et al., 1996)

Elliot et al. (1996) developed lists of team measures for AWACS crews. These lists cover *outcome measures* (overall measures of the degree to which teams accomplish mission goals), *indices of team SA* (degree to which members maintain current and accurate awareness and attentional focus), and *communications measures* (amount and type of communication). The measures in each list are specific to AWACS crews and may not be applicable in other domains. Nevertheless, the measures can serve as examples of the kinds of measures that can be used to assess the effectiveness of decision support and C2 team functioning.

TEAM MEASURES		
TEAM OUTCOME MEASURES	**SITUATION AWARENESS MEASURES**	**COMMUNICATION MEASURES**
Preservation of Assets	1. Total number of times assigned symbology of airborne track becomes uncorrelated	1. Total number of foot switch depressions
1. Friendly assets owned, not destroyed (overall measure)	2. Total number of times assigned symbology of downed track becomes uncorrelated	2. Total number of simulcast foot depressions
2. Friendly assets destroyed by hostile actions	3. Total number of times assigned symbology is assigned to wrong track	3. Total number of Net 2 foot switch depressions
3. Friendly aircraft shot down by hostile aircraft	4. Percept of SIM time at scale expansion 1	4. Total number of times a channel is switched off
4. Friendly aircraft lost to hostile SAM fire	5. Percept of SIM time at scale expansion 2	
5. Friendly assets jammed (soft kill)	6. Percept of SIM time at scale expansion 4	
Destruction of Enemy Assets	7. Percept of SIM time at scale expansion 8	
6. Hostile assets destroyed by friendly action (overall measure)	8. Percept of SIM time at scale expansion 16	
7. Hostile aircraft shot down by friendly aircraft	9. Percept of SIM time at scale expansion 32	
8. Hostile aircraft destroyed by friendly SAM	10. Total number of Initiate Switch actions button depressions	
9. Hostile aircraft lost to fuel out	11. Total number of Initiate Switch actions completed	
10. Hostile assets jammed (soft kill)	12. Total number of Initiate Switch actions correctly completed	

TEAM MEASURES		
TEAM OUTCOME MEASURES	**SITUATION AWARENESS MEASURES**	**COMMUNICATION MEASURES**
Efficiency/Effectiveness Process Measures 11. Kill ratio (sum of individual kill ratio, function of 3 and 5) 12. Air refuelings completed 13. Friendly aircraft lost to fuel depletion 14. Completed Assign/Defer switch actions 15. Total number of Commit Switch actions button depressions 16. Total number of Commit Switch actions completed 17. Total number of Commit Switch actions correctly completed 18. Total number of tactical bearing and range switch actions **Fratricide** 19. Friendly aircraft lost to friendly SAM 20. Friendly aircraft lost to friendly aircraft **Penetration** 21. Friendly penetration of hostile territory 22. Hostile penetration of friendly territory	13. Total number of Re-Initiate Switch actions button depressions 14. Total number of Re-Initiate Switch actions completed 15. Total number of Re-Initiate Switch actions correctly completed	

#510517